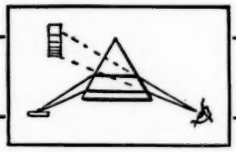
Journal of the OPTICAL SOCIETY *of* AMERICA

VOLUME 27
NO. 3



MARCH
1937

Quartz Mercury Arc with Electrodes of Low Work Function . . . . . WILLIAM T. ANDERSON, JR. AND LESTER F. BIRD 95

Graininess of Photographic Materials in Objective Absolute Measure . . . . . A. VAN KREVELD AND J. C. SCHEFFER 100

Capillary Chemistry and Fluorescent Light Microscopy . . . . . E. A. HAUSER AND C. J. FROSCH 110

Photometric Discrimination with a Diffuse Boundary . . . . . W. E. KNOWLES MIDDLETON 112

Photoelectric Method for the Determination of the Parameters of Elliptically Polarized Light . . . . . C. V. KENT AND JAMES LAWSON 117

On the Precision of Photometric Observations . . . . ALFRED H. HOLWAY 120

Further Search for the Ideal Color System I. A New Mechanico-Graphical Method . . . . . R. H. SINDEN 124

Brightness Meter . . . . . MATTHEW LUCKIESH AND A. H. TAYLOR 132

Hamilton's Characteristic Function and Bruns' Eiconal . . M. HERZBERGER 133

Hamilton's Characteristic Function and Bruns' Eiconal . . . J. L. SYNGE 138

Notice of Election of Officers for the Optical Society of America, Inc. . . . 145

Published monthly
at Prince and Lemon Sts., Lancaster, Pa.
by the
AMERICAN INSTITUTE OF PHYSICS
Incorporated
175 Fifth Avenue, New York, N. Y.
FOR AND UNDER THE EDITORIAL SUPERVISION OF
THE OPTICAL SOCIETY OF AMERICA

BOSTON UNIVERSITY
COLLEGE OF LIBERAL ARTS
LIBRARY

THE OPTICAL SOCIETY OF AMERICA

OFFICERS OF THE SOCIETY

A. C. HARDY, *President*
Massachusetts Institute of Technology
Cambridge, Massachusetts

R. C. GIBBS, *Vice President*
Department of Physics
Cornell University
Ithaca, New York

L. B. TUCKERMAN, *Secretary*
Bureau of Standards
Washington, D. C.

HENRY F. KURTZ, *Treasurer*
Bausch & Lomb Optical Company
Rochester, New York

The Board of Directors consists of the above officers; the Editor of the Journal; the Past President; and five elected members.

W. B. RAYTON, *Past President*
Bausch & Lomb Optical Company
Rochester, New York

F. K. RICHTMYER, *Editor*
Cornell University
Ithaca, New York

Elected Members

ARTHUR H. COMPTON
Ryerson Physical Laboratory
University of Chicago
Chicago, Illinois

K. S. GIBSON
Bureau of Standards
Washington, D. C.

G. R. HARRISON
Department of Physics
Massachusetts Institute of Technology
Cambridge, Massachusetts

HERBERT E. IVES
463 West Street
New York, New York

CARL W. KEUFFEL
Keuffel & Esser Company
Hoboken, New Jersey

ROCHESTER, NEW YORK SECTION OF THE OPTICAL SOCIETY OF AMERICA

L. A. DUBRIDGE, *President*
F. M. E. HOLMES, *Secretary*

B. NOYES, *Vice President*
C. A. MORRISON, *Treasurer*

Councilors: F. E. ALTMAN, F. C. FAIRBANKS, L. V. FOSTER, O. SANDVIK, C. C. NITCHIE (*ex officio*)

THE OPTICAL SOCIETY OF AMERICA was organized in 1916 "*to increase and diffuse the knowledge of optics in all its branches, pure and applied, to promote the mutual interests of investigators of optical problems, of designers, manufacturers and users of optical instruments and apparatus of all kinds and to encourage cooperation among them.*" It aims to serve the interests of all who are engaged in any branch of optics from fundamental research to the manufacture of optical goods.

MEMBERSHIP

Regular Members: "Any person who has contributed materially to the advancement of optics is eligible for regular membership."

Associate Members: "Any person who is interested in optics is eligible for associate membership."

Corporate Members: "Any corporation desirous of encouraging the work of the Society is eligible for corporation membership."

The Society cordially invites to membership all who are interested in any branch of optics, either in research, in instruction, in optical or illuminating engineering, in the manufacture and distribution of optical goods of all kinds or in physiological and medical optics.

ANNUAL DUES

Regular members, $7.50; Associate members, $5.00; Corporate members, $50.00. These dues include subscription to the Society's official journal, the *Journal of the Optical Society of America.*

MEETINGS

The Society holds two meetings per year: An annual meeting in the autumn and a midwinter meeting jointly with the American Physical Society.

For further information concerning the Optical Society of America address the Secretary, Dr. L. B. Tuckerman, National Bureau of Standards, Washington, D. C.

Journal of the
OPTICAL SOCIETY
of AMERICA

VOLUME 27 MARCH, 1937 NUMBER 3

A Quartz Mercury Arc with Electrodes of Low Work Function

WILLIAM T. ANDERSON, JR. AND LESTER F. BIRD[1]
Research Laboratory, Hanovia Chemical and Manufacturing Company, Newark, N. J.
(Received November 6, 1936)

THE quartz mercury arc has been designed, constructed and tested in an almost innumerable variety of forms. Until recently, only two of these forms have been found commercially practical. They are the familiar Cooper-Hewitt lamp and its modification, the quartz lamp by Kuch and Heraeus.[2]

These lamps were characterized by a liquid mercury cathode. They were essentially a direct current operated device, but have been extensively employed on alternating current in the form of a three electrode burner operated on a split secondary transformer and with a reactance in the neutral lead to the burner. This device is a rectifier.

Many early attempts were made to substitute solid cathodes for the liquid mercury. Most notable amongst these were the lamps[3] made by Konen and his associates at the University of Münster, Germany, during their investigations of gas spectra. They made both glass and quartz lamps, the latter in order that they might study the ultraviolet emission of various gases and metal vapors, and because the quartz glass permitted higher current inputs.

Konen was familiar with the work of Wehnelt[4] in the production of oxide cathodes of barium and calcium. He therefore used for his cathodes a small platinum foil 25 sq. millimeters in area and with a thickness of 0.1 millimeter. The foil, used for both the cathode and the anode, was covered with a thin layer of calcium and barium oxides. The arc tube in the quartz burner was 66 millimeters long and 2 millimeters bore. Gases and metal vapors were admitted to the lamp at pressures suitable for operation of the discharge. For argon this was about 10 millimeters. The lamp was operated on direct current supplied by an accumulator which had a no-load voltage of 440 volts.

The discharge in the tube could be started by rubbing the arc tube with a handkerchief, a piece of hard rubber, or even the hand. The lamp could also be started by means of a small induction coil. The initial discharge was a glow in the gas, which changed almost immediately into an arc discharge between the cathode and anode. The platinum oxide coated cathode was warmed by the discharge until it became incandescent. The cathode drop became only a few volts, and the arc voltage was 100 volts or less. The anode was strongly heated and became incandescent. The negative glow discharge disappeared. Current inputs as high as 1.5 amperes were used.

[1] Contributed paper No. 67 read before Optical Society, New York, October, 1936. Contribution from Research Laboratory, Hanovia Chemical and Manufacturing Company.

[2] Kuch, Physik Zeits. **6**, 438 (1905); Ann. d. Physik **20**, 563 (1906).

[3] Konen and Jungjohann, Verhandl. deutsch. Phys. Gesl. **3**, 145 (1910).

[4] Wehnelt, Ann. d. Physik **14**, 425 (1904).

BOSTON UNIVERSITY
COLLEGE OF LIBERAL ARTS
LIBRARY

Konen recommended that his lamp be made into a practical metal vapor arc lamp. This was not accomplished because the electrode structure was not sufficiently durable for a commercial lamp.

In 1919, Henri George[5] demonstrated in Paris a quartz mercury arc lamp which was a modification of the Konen lamp. George substituted ductile tungsten, which had become available during the time interval between Konen's experiments in 1909 and 1919, for the activated platinum. He appears to have thought that employment of the much higher melting metal would overcome the difficulties of the Konen lamp. In this he was mistaken.

The greatest stress on cathodes occurs at the start of the discharge. The electron emission from a cathode is a function of the temperature. Even under the most favorable conditions, the emission is relatively very small when the cathode is at or near room temperature. Another factor which requires consideration is that the current in a vapor arc is greater at the start than after the equilibrium state has been attained. It becomes, therefore, a matter of necessity that the work function of the cathode be below 3, and that the cathode be proportioned in such a way that very rapid warming with its greatly increased emission results. Finally the cathode temperature must not exceed a temperature above which the cathode is destroyed.

Konen's electrodes partially met the above requirements. They had a work function of about 3. They deteriorated rapidly, however, because they were not properly proportioned. The tungsten electrodes used by George could withstand the temperature, but sputtered badly. They had a work function of 4.4.[6] The George lamp had an operable life of relatively few hours and the quartz envelope became rapidly blackened. The life of the lamp was greatly dependent upon the frequency of lighting.

In the early 1920's, Davisson and Germer,[7] and Spanner[8] made careful studies of the emissivity of a great many oxides. They were familiar with the Konen lamp, and they appreciated its shortcomings. As a result of these and later researches, Spanner and Doering,[9] Meyer,[10] and Germer[11] developed a practical lamp in which the oxides and platinum foils of Konen, and the tungsten electrodes of George were replaced by spirals of wire coated with alkaline earth oxides fused with alumina or zirconia, the latter serving to make the activation coating more resistant to the destructive action of the discharge. Their electrodes, like those of Konen, were heated by the direct action of the discharge after starting with cold electrodes. This design provided the rapid warming and durability which was missing in the Konen lamp. Their method of construction resulted in a commercially practical light source, which in a modified form, has been successfully marketed in Europe.

In the United States, independent investigations of quartz mercury arc lamps made in this laboratory have resulted in the construction of lamps with solid self-heating activated electrodes in which rapid warming was procured and destruction was prevented by a method of automatic temperature control. It had been noted that some oxides, notably those of aluminum, zinc and zirconium, were very poor conductors for heat when at or near room temperature. When heated, these oxides became increasingly conductive as the temperature increased, until at incandescence, they became very good conductors.

The thought was suggested that suitable electrodes might be constructed in the form of a laminated structure comprising pieces of metal separated by these oxides. The outermost metal surface would bear the Wehnelt thermionic materials, barium and calcium oxides fused in position. The different metal layers would be connected mechanically and electrically by a small electric weld.

In operation, such an electrode would function as follows: The discharge would strike the outermost metal surface bearing the fused activated materials. This surface would warm rapidly because it was of small area and insulated ther-

[5] Henri George, U. S. Patent No. 1,361,710, December 7, 1920.
[6] *International Critical Tables* (McGraw-Hill, 1929), Vol. 6, p. 53.
[7] Davission and Germer, Phys. Rev. **20**, 300 (1922); ibid. **24**, 666 (1924).
[8] Spanner, Ann. d. Physik **75**, 609 (1924).
[9] Spanner and Doering, Brit. Patent No. 397,162, Aug. 8, 1933.
[10] Meyer and Spanner, Brit. Patent No. 274,027, Oct. 20, 1927.
[11] Germer, French Patent No. 707,198, April 13, 1931.

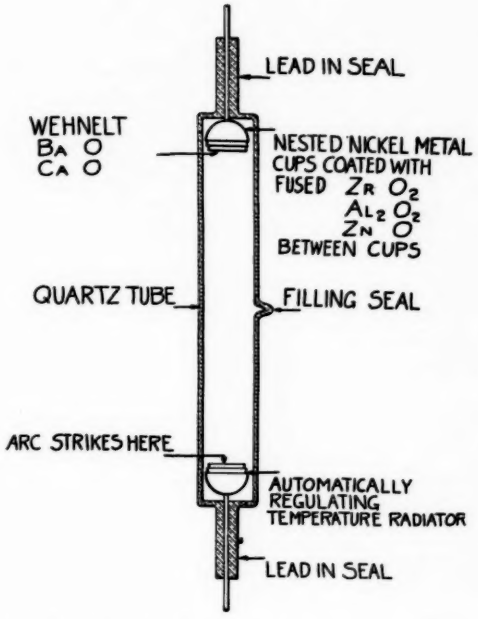


FIG. 1. Quartz burner showing the electrode arrangement.

mally. Before a destructive temperature could be attained, the layers of insulating oxides would become conductive and warm the next layer in the laminated structure. This action would be repeated as many times as required to attain the desired stable operating temperature.

Lamps and electrodes embodying these ideas were constructed. In the earliest forms, the final temperature was controlled by conducting the heat from the electrodes to the walls of the vessel.[12] This procedure was not so satisfactory because the final temperature of the electrodes became too dependent upon the surrounding temperature of the lamp. A more self-contained structure was devised[13] in which the heat was radiated directly from the electrode itself. This consisted of a number of nested metal cups connected mechanically and electrically by a weld, and separated from one another by the oxide coatings used in the automatic thermal regulator. The cups were supported by a single support conductive lead.

Fig. 1 illustrates a typical burner constructed with these automatically thermally regulated electrodes. The materials comprising the metallic supports and radiators, in this instance the cups, have been found to be unimportant provided that they do not melt at the operating temperature of the electrode. For quartz mercury arcs, nickel, platinum, molybdenum, palladium, tantalum and tungsten are satisfactory.

The gas filling in the lamp is the usual one employed by Konen, George, Spanner *et al.*, namely, argon at about one centimeter pressure, and sufficient mercury to provide a vapor filling at the operable wattage.

The operating characteristics are considerably dependent upon the dimensioning of the quartz vessel. The 200-watt burner illustrated which is operable on 220 volts a.c. gives the best results when the arc length is about 130 millimeters and the tube diameter is about 16 millimeters. Like all arcs, a ballast resistance or a reactance must be used in series with the burner to limit the current.

The lowest practical striking potential for argon is about 180 volts. It is therefore either necessary to have a supply voltage higher than this, or to use a shocking device such as a condenser or inductance. By the employment of such a device, it is even possible to start and to operate properly proportioned burners upon 110 volt a.c. supply. Starting can be further assisted by the employment of a grounding device which collects electrostatic charges from the walls of the vessel. A very convenient form of this device is a stripe on the outside of the tube made by painting on the tube with a ceramic gold or platinum paint. These materials are conducting.[14]

Henri George in his United States patent[5] gives an excellent description of the functioning of these types of lamps which we believe to be worthy of quotation. "The functioning is as follows: On the closing of the interrupter, the tube is immediately lighted up as a luminescence tube, the arc passing through the inert gas atmosphere. The mercury is at once vaporized; the intensity increases and passes from the order of luminescence to that of the arc in mercury vapor. The internal pressure exceeds one kilogram per

[12] Anderson, U. S. Patent No. 1,902,936, March 28, 1933.
[13] Anderson and Bird, U. S. Patent No. 2,006,081, June 25, 1935.
[14] Spanner, U. S. Patent No. 1,860,210, May 24, 1932.

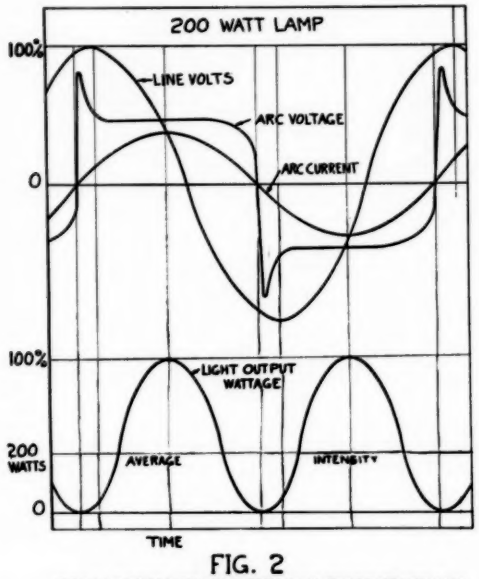


FIG. 2
OSCILLOGRAM OF WAVE FORM AND PHASE RELATIONS OF ELECTRICAL CHARACTERISTICS AND LIGHT OUTPUT.

square centimeter, and the arc assumes the form of a very brilliant cord."

The Konen and George lamps both functioned in this manner, but only for a short time. In the Konen lamp the electrodes became overheated and melted. In the George lamp the electrodes sputtered and rapidly blackened the tube. In both instances, the burners failed after only a few hours of use. In the lamps described in this paper, the electrodes do not overheat and melt; neither do they sputter. As long as the electrode coatings remain intact (it does eventually wear out), the burner will continue to function. As a result, these burners will operate under proper conditions for about two thousand hours, and can be lighted many hundreds of times without failure. The lamp appears to be a practical source for ultraviolet radiation.

Fig. 2 shows an oscillogram of the wave form and phase relations of the electrical characteristics of a quartz burner when operated on 60-cycle alternating current. The curve for the arc voltage indicates that the burner actually is extinguished each time that the frequency changes. The peak in the burner voltage curve is the ignition voltage. As might be expected, the arc current also passes through zero at the change of cycle. In the lower portion of the figure it is indicated that the light output is a function of the burner wattage. This is generally characteristic of arc lamps. The light output follows the arc current, reaching a maximum when the current is at a peak, and, of course, becoming zero when the lamp is extinguished between cycles. The lamp is therefore stroboscopic in nature, but in appearance to the eye, there is a continuous output of light.

Fig. 3 shows the volt, watt and ampere characteristics of the burner operated on alternating current from the time of starting until the equilibrium steady state has been reached. As in the lamp by George, all or most of the mercury has been volatilized when this condition has been reached. Limitation of the quantity of mercury is desirable because it makes the arc less subject to variations resulting from the cooling effects of air drafts. The tubes which contain much less mercury than the older forms of quartz mercury arc lamps are for this reason much less likely to be broken during shipment.

Since the electrodes used in this lamp are oxide electrodes, there is always present in the

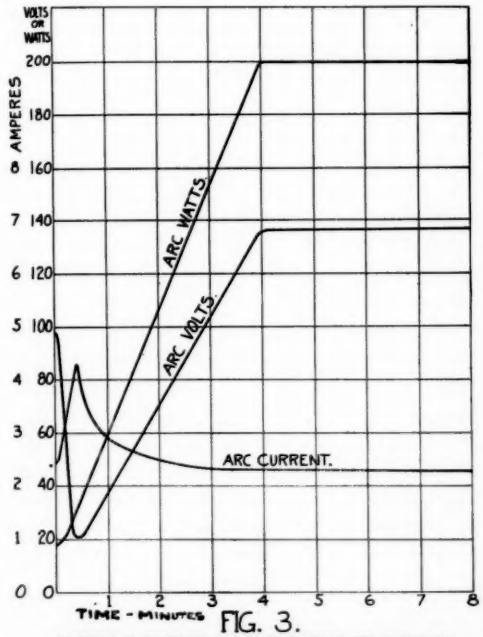


FIG. 3.
TYPICAL ELECTRICAL CHARACTERISTICS OF 200 WATT 220 VOLT LAMP OPERATED WITH TRANSFORMER ON 110 VOLTS, SERIES INDUCTANCE 0.119 h.

tube an exceedingly low pressure of oxygen gas. If this oxygen pressure became too high, it would interfere with the starting of the discharge. This is prevented by the mercury vapor itself which serves as a getter to maintain this oxygen pressure at a very low value. The quantity of mercury consumed in this manner is so small that it is of no consequence to the operation of the lamp.

These lamps are hot high pressure quartz mercury arcs. The mercury vapor pressures used have approximated atmospheric. However, it seems probable that higher pressures might be used if desired and if the tube were properly proportioned. The spectrum of the light from the lamps and its spectral energy distribution are in every respect similar to all other hot high pressure quartz mercury arcs at an equivalent mercury vapor pressure. This spectrum and the energy curve are so familiar to those acquainted with the art that they have not been included in this paper.

Watt for watt, this burner makes possible a more efficient conversion of electrical energy into light than can be obtained in the usual forms of quartz mercury arcs because of lower electrode losses. The electrical operating equipment required for alternating current operation is simpler in form than that usually employed, and electrical losses in it are likewise considerably less. The arc operates equally well in any position, and can be started by the snap of a switch.

This new form of the quartz mercury arc offers many advantages over the older forms, and should be a useful addition in the field of ultraviolet irradiations.

Graininess of Photographic Materials in Objective Absolute Measure

A. VAN KREVELD AND J. C. SCHEFFER
Netherlands Foundation for Photographic and Cinematographic Research

(Received September 22, 1936)

List of symbols used in this and former papers.[1-3]

I: intensity of the exposure.
D: logarithmic density caused by the exposure.
γ_d: the differential coefficient $dD/d \log I$ at the density D.
γ_t: the total gradation, i.e., $(D_2-D_1)/(\log I_2-\log I_1)$ where 1 and 2 denote the lower and upper limit of the useful part of the density curve.
ΔD: the average fluctuation of the density, caused by the graininess.
ΔD_a: the apparent average fluctuation of the density, to be corrected for the inertia of the galvanometer.
ΔD_r: the true average fluctuation of the density, derived from ΔD_a by correcting for the inertia of the galvanometer.
ΔK: the absolute graininess, i.e., $\Delta D_r \times$(observed area)$^{\frac{1}{2}}$.
$\Delta \log I$: $=\Delta I/I=\Delta D/\gamma_d$: the average fluctuation of the intensity derived from ΔD.
$\Delta_K \log I$: $=\Delta_K I/I=\Delta K/\gamma_d$: the average fluctuation of the intensity derived from ΔK.
G: the average distance between the points of intersection of the line representing average density with the trace representing variations in density (this magnitude is canceled now, because it is merely a function of ΔD and of the scanned area).
L: the turbidity = the distance in the emulsion, where the intensity of the scattered light is half the intensity of the incident light.
Q: the micro-quality of the emulsion for a given application.

INTRODUCTION

IN former papers[1-3] one of us has described a method for the objective measuring of the graininess of photographic materials. This method is based on the fluctuations of the density caused by the irregular distribution of the grains. The mean value of these fluctuations denoted by ΔD provides a measure for the graininess. The measuring of ΔD is performed by moving the negative through two parallel light beams, each of which strikes on a barrier-layer cell. The graininess of the plate causes fluctuations in the intensity of the transmitted light, which cause in their turn fluctuations in the current of the cells. The cells are connected in such a way, that the difference of their currents is measured. The mean of the absolute value of this difference is proportional to ΔD. For a full description see reference 3.

One objection has been raised against this method: the results depend on the width of the light beams and on the period and degree of damping of the recording galvanometer. In this paper we shall show how all arbitrary properties of the apparatus can be eliminated so that an absolute measure is obtained, which we shall call the absolute graininess ΔK.

Further we shall report on a number of practical measurements with the graininess-meter described.

THE INFLUENCE OF THE GALVANOMETER

The galvanometer, which is connected to the barrier-layer cells, does not follow the fluctuations of the current exactly. The distribution of the fluctuations being at random, all frequencies will be present, among them very slow as well as very rapid fluctuations. Now, however quick the galvanometer may be, there will always be some fluctuations which it cannot follow. The complete theory of these problems has been given in a paper of Ornstein and van Kreveld.[4] From this paper we cite the following: "We suppose the galvanometer being aperiodically damped (this is true of the galvanometer used in the graininess-meter). Further we suppose the graininess being measured by means of a rectangular slit (e.g. the slit of a microphotometer); then the following relation holds good:

$$\Delta D_a=\Delta D_r\left(\frac{(V\tau+3)e^{-V\tau}+2V\tau-3}{2V\tau}\right)^{\frac{1}{2}}, \quad (1)$$

[1] A. van Kreveld, Phot. J. **74**, 590–596 (1934).
[2] A. van Kreveld, Communication IX. Int. Phot. Congress Paris, pp. 263–267 (1935).
[3] A. van Kreveld, J.O.S.A. **26**, 170–176 (1936).
[4] L. S. Ornstein and A. v. Kreveld, Physica **3**, 815–822 (1936).

V being the constant from the well-known differential equation for the aperiodic galvanometer:

$$d^2x/dt^2+2Vdx/dt+V^2x=F/K$$

and τ being the time in which the negative moves over a distance equal to the width of the image of the slit."

In the case of the graininess-meter the image on the negative has not the simple rectangular form; it consists of two concentric circles (see Fig. 1).

The outer part a is focused on one barrier-layer cell, the inner part b on the other cell, so that the fluctuations of the difference $a-b$ are measured. In this case the relation (1) must be replaced by:

$$\Delta D_a=\frac{\Delta D_r}{2}\left[\frac{1}{V\tau}\{e^{-V\tau\sqrt{2}}(2V\tau+3\sqrt{2})+e^{-\frac{1}{2}(\sqrt{2}+1)V\tau}\times\{-(2\sqrt{2}+2)V\tau-12\}+e^{-V\tau}(4V\tau+12)+e^{-\frac{1}{2}(\sqrt{2}-1)V\tau}\{(2\sqrt{2}-2)V\tau+12\}+4V\tau-12-3\sqrt{2}\right]^{\frac{1}{2}}. \quad (2)$$

This formula can be proved in the same way as relation (1). The proof requires, however, lengthy calculations. It will be given in the Utrecht thesis of one of us (Sch.).

Henceforth we shall reduce all measurements by the formula (2) without further reference. In this way *the influence of the galvanometer is eliminated completely.*

TABLE I. *Plates measured with different diaphragms.*

No.	BRAND	DEVELOPER (SEE TABLE IV)
1	Ilford Special Rapid	Metolborax A.
3	Gevaert Contrast	Hydroquinone B.
4	Gevaert Supercontrast	Paraphenylenediamine D.

TABLE II. *Data about the diaphragms.*

No.	DIAMETER IN MM	AREA OF THE IMAGE ON THE NEGATIVE IN μ^2	τ IN SECONDS	CORRECTING-FACTOR ACCORDING TO (2)	CORRECTING-FACTOR $\times$AREA$^{\frac{1}{2}}$ IN μ
I	2.78	22220	0.295	1.50	224
II	2.06	12200	0.22	1.655	183
III	1.38	5480	0.147	1.92	142
IV	1.04	3110	0.111	2.14	119.5

FIG. 1. The shape of the image formed in the graininess-meter.

THE INFLUENCE OF THE MEASURED AREA

In the second place we have investigated the dependence of ΔD on the width of the slit, or more correctly: on the size of the circular image (see Fig. 1).

For that purpose we have measured the graininess of a number of plates with four diaphragms of different sizes. The data about the plates and diaphragms are given in the Tables I and II.

The diaphragms consist of brass plates with circular holes in them. In the graininess-meter an image of the holes is focused on the negative with a reduction of $\times 11.7$.

The results of the measurements (without correcting for the galvanometer) are given in the left hand side of Table III. Now, if the distribution of the graininess is normal, we may expect that the true average fluctuation of the density ΔD_r is inversely proportional to the square root of the measured area. In order to test this supposition we have calculated the quantity ΔK defined by the relation

$$\Delta K=\Delta D_r\times\text{area}^{\frac{1}{2}}$$
$$=\Delta D_a\times\text{correction factor}\times\text{area}^{\frac{1}{2}}.$$

The results are given in the right hand side of Table III. From Table III it appears that for all negatives ΔK *is perfectly independent of the size of the diaphragm.* So it may be used as an absolute measure for the practical graininess. As ΔK has the dimension of length we shall express it in microns.

THE MEANING OF ΔK

As all results of our graininess measurements henceforth will be expressed in ΔK, it may be useful to explain the meaning of this quantity in some detail. We begin with an example. Suppose that for a given photographic emulsion

in combination with a given developer at a given density the average fluctuation ΔD_r in the graininess-meter proves to be 0.0065. Further we suppose that the circular image (see Fig. 1) has a diameter of 0.2 mm. Then each part of the image has an area of $\frac{1}{2}\times\pi/4\times200^2\mu^2=15{,}700\mu^2$; hence $\Delta K=0.0065\times(15{,}700)^{\frac{1}{2}}=0.815\mu$. Now we can calculate by means of ΔK the density fluctuations for any application. We give three examples:

1. Sound on film: the size of the image of the slit on the sound-track is $2500\mu\times10\mu=25{,}000\mu^2$. The average fluctuation of the density is $\Delta K/\text{area}^{\frac{1}{2}}=0.815/(25{,}000)^{\frac{1}{2}}=0.0051$.

2. For the measuring of band-spectra with the microphotometer of Moll a very narrow slit is used. The image of this slit on the negative has a size of $400\mu\times2\mu=800\mu^2$. Hence $\Delta D=0.815/(800)^{\frac{1}{2}}=0.029$.

3. Enlargements from miniature negatives are made with lenses with a limited resolving power, e.g. 0.01 mm. Then the area which must be considered for the graininess is $(\pi/4)\times100\mu^2=78.5\mu^2$ and $\Delta D=0.815/(78.5)^{\frac{1}{2}}=0.092$.

These examples show that the calculation of the practical graininess for various applications is very simple indeed. We should not, however, lose sight of the limits of this method. The inverse proportionality of ΔD with the square root of the area will naturally only hold good, when the area is large compared with the size of each individual grain. This condition is equivalent to: $\Delta D\ll1$.

In practice this condition is always satisfied, the resolving power of the usual optical systems being low compared with the size of the silver grain. So the lower limit for the validity of our calculations can safely be estimated to be *circa* $100\mu^2$.

There exists on the other side an upper limit for the accuracy of the formula; this becomes clear if we keep in mind that in the case of very large areas ΔD becomes so small, that it is negligible in comparison with the macroscopic inhomogeneities of the emulsion caused by the coating process. The latter are quite irregular and cannot be expected to satisfy any statistical law. It is, however, difficult to give a sufficient estimation of the magnitude of the upper limit. In any case it must be large compared with the size of the image of our widest diaphragm ($22{,}220\mu^2$), the latter not showing the slightest deviation from the Gaussian law. Our calculations can, therefore, safely be applied between

TABLE III. *ΔD_S measured and ΔK computed from the measurements.*

Density	ΔD_S				$\Delta K=\Delta D_S\times\text{CORR.}\times\text{AREA}^{\frac{1}{2}}$ IN μ			
	Diaphragms				Diaphragms			
PLATE 1 A.	I	II	III	IV	I	II	III	IV
0.08	0.0013	0.0017	0.0022	0.00285	0.29	0.31	0.31	0.34
0.17	18	255	295	37	0.40	0.465	0.42	0.44
0.35	245	305	38	48	0.55	0.56	0.54	0.575
0.55	29	37	51	60	0.65	0.675	0.725	0.715
0.78	315	42	53	555	0.705	0.77	0.75	0.665
1.00	35	425	48	56	0.785	0.78	0.68	0.67
PLATE 3 B.								
0.07	0.00105	0.00125	0.00145	0.0019	0.235	0.23	0.205	0.225
0.14	155	175	215	24	0.345	0.32	0.305	0.285
0.27	205	235	335	425	0.46	0.43	0.475	0.505
0.50	265	29	44	51	0.595	0.53	0.625	0.61
0.88	36	43	55	68	0.805	0.785	0.78	0.81
PLATE 4 D.								
0.08	0.0008	0.00095	0.0011	0.00135	0.18	0.175	0.155	0.16
0.19	11	14	165	21	0.245	0.255	0.235	0.25
0.355	15	19	24	29	0.335	0.35	0.34	0.345
0.56	175	21	30	33	0.39	0.385	0.425	0.395
0.79	19	23	26	34	0.425	0.42	0.37	0.405
1.085	21	25	30	395	0.47	0.46	0.425	0.47

10^2 and $10^5\mu^2$, and for this range the meaning of ΔK can be expressed in the following way: *The mean deviation in density of a square with sides of C microns is equal to* $\Delta K/C$ (ΔK being expressed in μ).

Besides ΔK we shall consider two other quantities important for the study of the practical graininess, namely $\Delta K/\gamma_d$ and $\Delta K/\gamma_t$ (see list of symbols). $\Delta K/\gamma_d$ is equal to the fluctuation $\Delta K \log I$ of the logarithm of the intensity, implied in the density fluctuation ΔK. It qualifies the graininess for the following applications (for full description see reference 1): measurement of distances (positions of stars, wave-lengths of spectral lines) and photographic photometry.

On the other hand the magnitude $\Delta K/\gamma_t$ is essential for indicating the graininess for other applications, e.g. enlargements of miniature negatives and sound recording.

Correction for the Nonlinearity of the Rectifier

Measuring the graininess at high densities, we observed a rapid decrease of ΔD with increasing density, which could not be expected *a priori*. This phenomenon which was already mentioned in a former paper[3] appeared to be equivalent to a nonlinearity of the optical rectifier for very small readings, caused by a lack of definition of the image of the moving rectangle. This is unavoidable, the diffraction of the light on the small mirror of the first galvanometer being the cause.

In order to find the correction for the nonlinearity we proceeded as follows: If the rectangular image were perfectly sharp and had a width equal to the width of the strip of metal which checks it at zero-deflection, the characteristic of the rectifier would be given by the line a of Fig. 2. If, however, the perfectly sharp image were $2r$ millimeters narrower or wider than the strip of metal, the characteristic would be given by the line b, c, respectively. In both cases there would be a threshold deflection of r millimeters. Hence the case a provides the highest sensitivity.

Actually the edges of the image are not perfectly sharp and the real characteristics are given by the corresponding curves a_1, b_1, c_1 of Fig. 3. The most favorable case is a_1, so we have adjusted the magnification of the image in such a way that this case is realized.

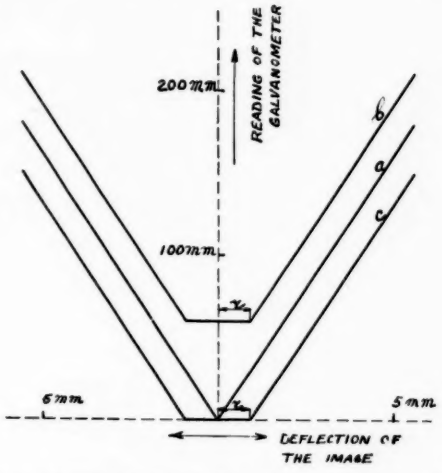


FIG. 2. Characteristics of the rectifier in the case of a perfectly sharp image.

The characteristics a, b, c, a_1, b_1 and c_1 apply to constant deflections. For graininess-measurements, however, the deflection is fluctuating and has a Gaussian distribution (see reference 3). So we cannot use the curve a_1, but have to take a somewhat different curve a_2, calculated from a_1 by superposing a Gaussian distribution on it.

On each day that measurements were performed we determined the curve a_1 experimentally and calculated the curve a_2 from it. This curve appeared to be nearly constant. All values for the graininess given in this paper are corrected according to the curve a_2.

Practical Applications

In order to determine the influence of various practical conditions for the measurement of the graininess we have investigated the value of ΔK for approximately 400 photographic images.

All of the negatives were exposed to the light of a 6 volt-6 amp. lamp, combined with a blue glass, transmitting a narrow band at 4600A. Each image was a circle of constant density with a diameter of 15 mm. On every plate we took successively 8 to 12 images of different density. The exposure time was 30 seconds in all cases and the intensity of the light was changed by varying the current of the lamp.

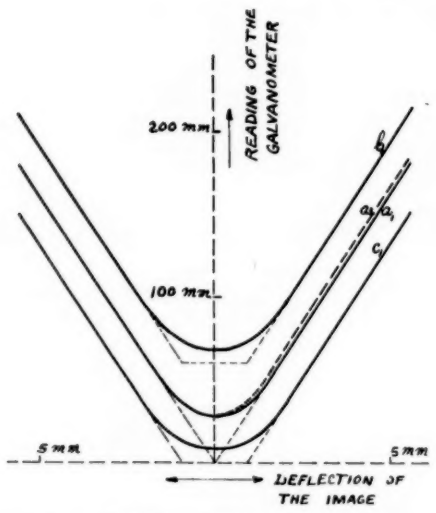


FIG. 3. Real characteristics of the rectifier.

The calibration of the lamp was performed by the usual method of the physical laboratory of Utrecht.[4a] All plates were developed separately in a tray at a temperature of 18°C. The fixing bath consisted of: 1000 cm³ water, 150 g sodium thiosulfate, 15 g potassium metabisulfite. After fixation the negatives were washed for an hour in running water, rinsed in distilled water and dried in a special box free from dust. After that the graininess of each image was measured four times with diaphragm II. The densities were also measured with the graininess-meter. This procedure has the advantage that the density and the graininess are measured with equal light-distribution.

In the first place we studied the influence of various developers on fine grained and coarse grained emulsions. We used for this purpose the following brands:

1. Ilford Special Rapid	400 H & D.
2. Ilford Ordinary	70 H & D.
3. Gevaert Contrast	500 H & D.
4. Gevaert Supercontrast	100 H & D.

The speed numbers indicated are those given by the manufacturers. The sensitivity of the Gevaert plates turned out to be much lower. The successive numbers 1, 2, 3 and 4 given above correspond with decreasing true sensitivity and decreasing size of the grains. In combination with these four emulsions we used eight developers *A–H* (described in Table IV), selected as different as possible. The results of the $4\times8=32$ combinations of emulsion and developer are given in Figs. 4–15. Figs. 4–7 give the density-curves, Figs. 8–11 give ΔK plotted against the density, Figs. 12–15 give $\Delta K/\gamma_d$ plotted against the density (fog always included).

The following remarks can be made about the general shape of the graininess-curves: The quantity ΔK always increases uniformly with increasing density and reaches its maximum at a density of about 1.0. This maximum is very flat and is followed by a slow decrease.

The shape of the $\Delta K/\gamma_d$ curves is much more complicated. At very low densities γ_d tends to zero, while ΔK always keeps a finite value caused by the fog, i.e. all $\Delta K/\gamma_d$ curves touch the vertical axis asymptotically. So $\Delta K/\gamma_d$ decreases with increasing density very rapidly at first, reaches in most cases a minimum, after that a maximum and finally decreases very slowly. The minimum and maximum are sometimes absent (e.g. the curves 1*H*, 3*A*).

The complicated shape of the $\Delta K/\gamma_d$ curve has no essential meaning. It arises from the fact

TABLE IV. *Formulae for the developers.*

A. *Metolborax:*[5] Metol, 8.6 g; borax, 1.9 g; Na_2CO_3, 5.3 g; $Na_2SO_3.7H_2O$, 75 g; KBr, 0.5 g; water up to 1000 cm³.
Time of development 6 minutes.
B. *Hydroquinone:* Hydroquinone 3 g; $K_2S_2O_5$, 3 g; KOH, 12 g; KBr, 3 g; water up to 1000 cm³.
Time of development 6 minutes.
C. *Ferrous oxalate:*[6]
I water 100 cm³; $K_2C_2O_4$, 25 g.
II water 100 cm³; $FeSO_4$, 30 g; citric acid 1 g.
III water 100 cm³; KBr, 1 g.
The developer consists of 80 cm³ of I, 20 cm³ of II and two drops of III.
Time of development 6 minutes.
D. *Paraphenylenediamine:*[7]
Paraphenylenediamine 10 g; $Na_2SO_3.7H_2O$, 120 g; Na_3PO_4, 2 g; KBr, 1 g; water up to 1000 cm³.
Time of development one hour.
E. 100 cm³ of *A* diluted with 100 cm³ water.
Time of development 6 minutes.
F. 100 cm³ of *B* with 50 cm³ water.
Time of development 6 minutes.
G. 100 cm³ of *C* with 100 cm³ water.
Time of development 6 minutes.
H. Same as *D*, but time of development 30 minutes.

[4a] L. S. Ornstein, E. F. M. van der Held and D. Vermeulen, J.O.S.A. **20**, 573 (1930).

[5] W. Reinders and M. Beukers, Bericht VIII. Int. Phot. Congress, Dresden (1931), p. 177.

[6] G. Eberhard, Publ. Potsdam **26**, No. 84, *H.* 1 (1926).

[7] Lumière-Seyewetz according to J. M. Eder: Rezepte, Tabellen; 14–15. Aufl. p. 76.

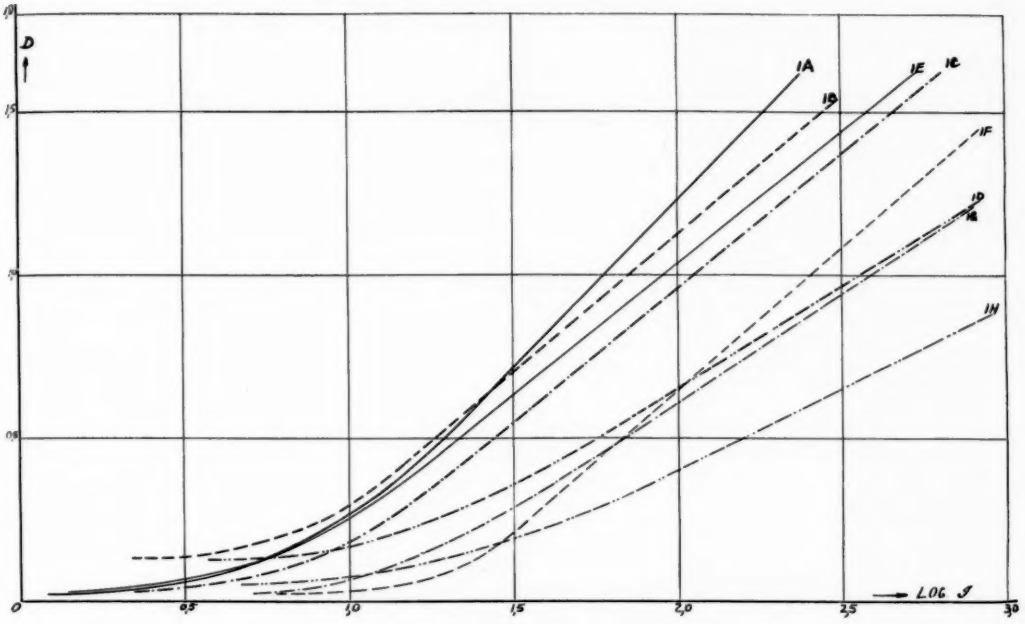


FIG. 4. Density curves of Ilford Special Rapid.

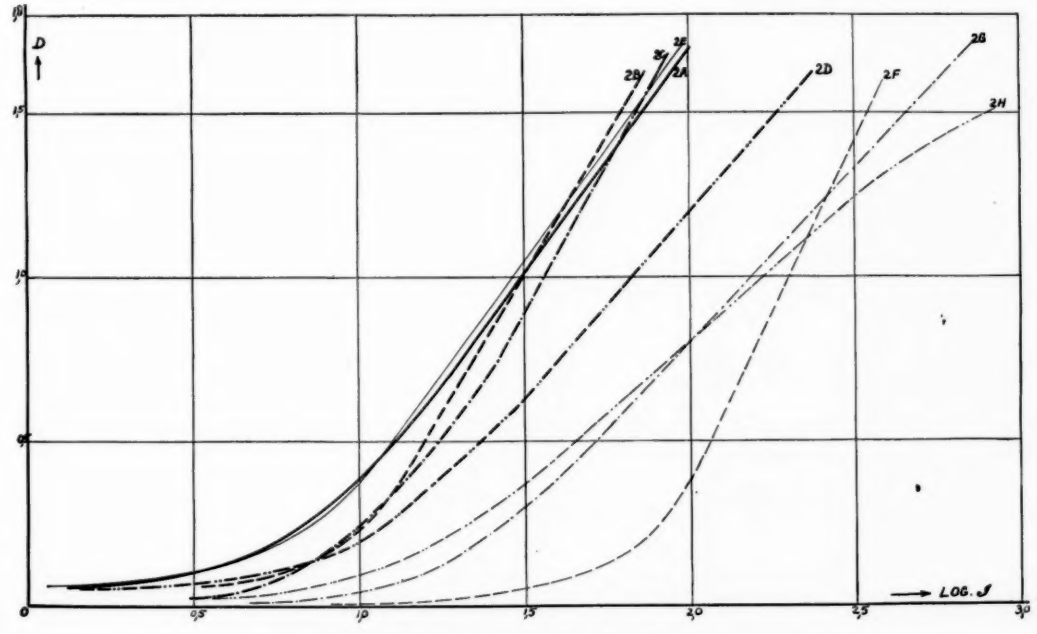


FIG. 5. Density curves of Ilford Ordinary.

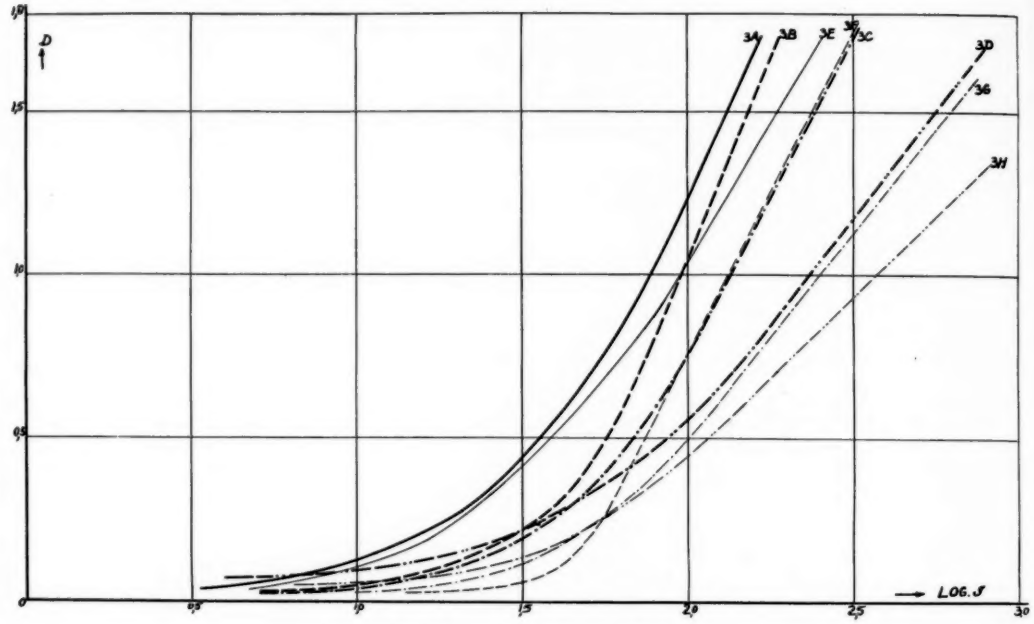


FIG. 6. Density curves of Gevaert Contrast.

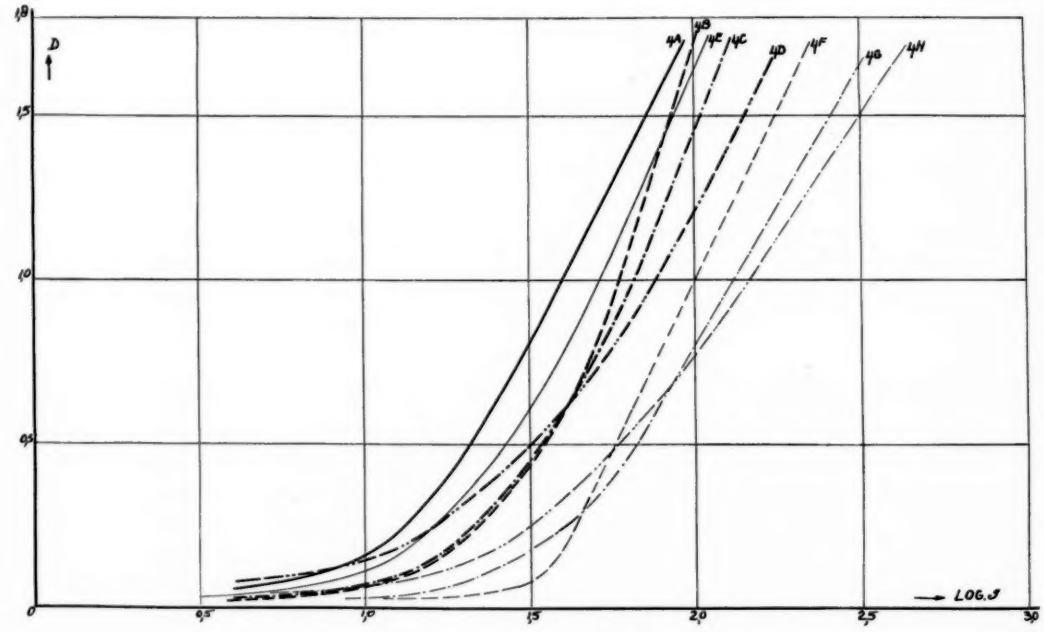


FIG. 7. Density curves of Gevaert Supercontrast.

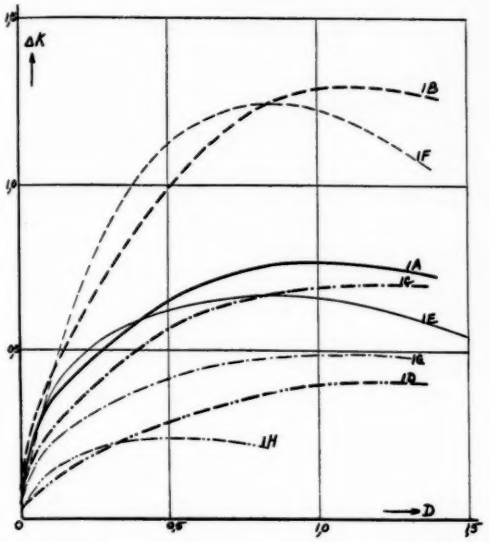


FIG. 8. Absolute graininess curves of Ilford Special Rapid.

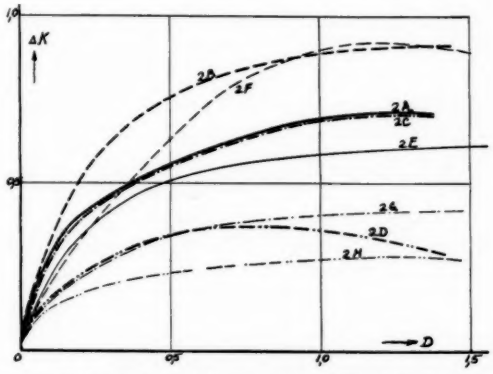


FIG. 9. Absolute graininess curves of Ilford Ordinary.

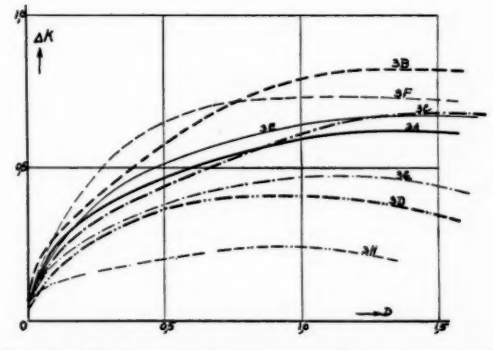


FIG. 10. Absolute graininess curves of Gevaert Contrast.

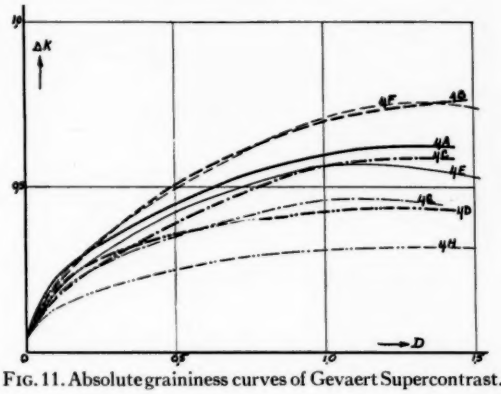


FIG. 11. Absolute graininess curves of Gevaert Supercontrast.

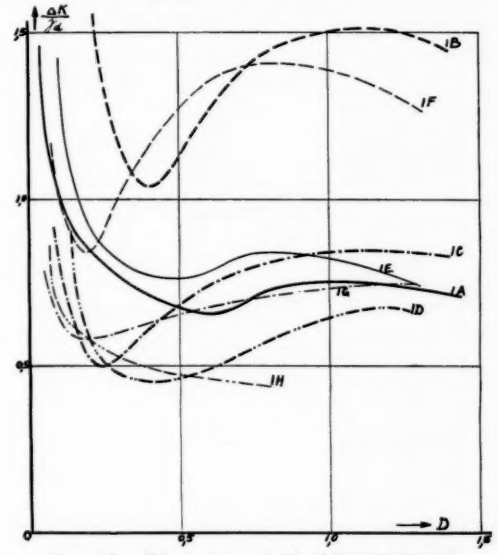


FIG. 12. $\Delta K/\gamma_d$ curves of Ilford Special Rapid.

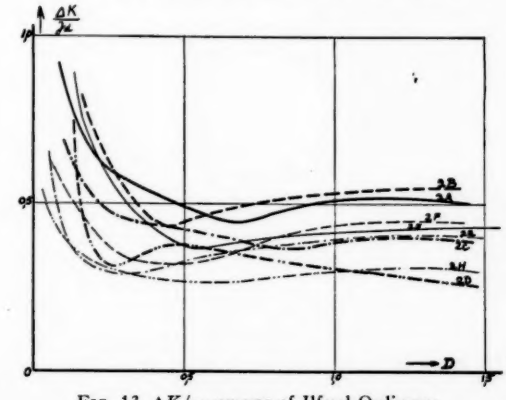


FIG. 13. $\Delta K/\gamma_d$ curves of Ilford Ordinary.

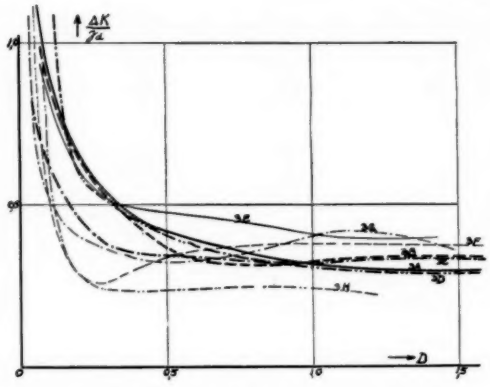


FIG. 14. $\Delta K/\gamma_d$ curves of Gevaert Contrast.

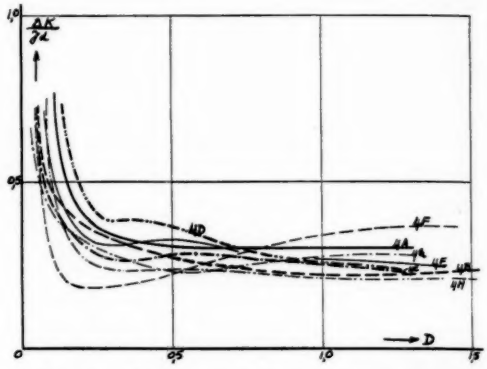


FIG. 15. $\Delta K/\gamma_d$ curves of Gevaert Supercontrast.

that $\Delta K/\gamma_d$ is the quotient of two functions, both having a maximum. The accidental mutual position of these maxima determines, whether the $\Delta K/\gamma_d$ curve will have a maximum and a minimum or will be quite monotonous. In any case the $\Delta K/\gamma_d$ curve has a fairly constant value in the most important range of densities and we may take this mean value as a rough numerical measure for the graininess of the combination emulsion-developer. These mean values are indicated in Table V. (Applicable to measurement of distance and photographic photometry.)

The mean values of $\Delta K_{max}/\gamma_t$ are given in Table VI (applicable to making of enlargements and sound-film).

The differences in the ratios between the Tables V and VI may be neglected; *for most practical cases the number* $\Delta K_{max}/\gamma_t$ *will be sufficient to characterize the graininess of the combination emulsion-developer.* The data of the figures and of Tables V and VI allow us to draw the following conclusions:

1. Our method is applicable with sufficient precision both to coarse grained and to fine grained emulsions.

2. It is not possible to decide conclusively which developer gives the lowest graininess, the quality of the developer depending to some extent on the emulsion, with which it is combined, e.g.: the metol-developer *A* and the hydroquinone-developer *B* cause equal graininess on emulsion 2, while the graininess which they cause on emulsion 1, differs by a factor of two. This example, however, is exceptional. In most cases the results are more regular and may be expressed as follows: the influence of the developer is most considerable on the coarse-grained emulsion 1 (and 2) and can be neglected for the fine-grained emulsion 4. This result is exactly, what might be expected.

In all cases paraphenylenediamine (D and H) gives the lowest graininess.

3. The results of the plates 1*B* and 1*F* are very remarkable. The hydroquinone-developer causes on this coarse grained emulsion an extremely high graininess, which proved to be quite reproducible. The finer grained emulsions do not show this phenomenon. We intend therefore to study it on some other coarse grained emulsions. Possibly it is caused by the high *p*H of our formulae *B* and *F*.

FURTHER MEASUREMENTS

1. In order to complete our investigation we have measured the graininess of two orthochromatic superspeed rollfilms, developed according to formula *A* (metol). The graininess differs not much from that of the Ilford Special

TABLE V. *Mean values of* $\Delta K/\gamma_d$ *between densities of 0.1 and 1.5 (fog subtracted).*

	DEVELOPER							
PLATE	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*
1	0.71	1.38	0.75	0.55	0.80	1.28	0.68	0.48
2	0.50	0.51	0.40	0.32	0.42	0.39	0.37	0.29
3	0.36	0.36	0.34	0.36	0.43	0.35	0.36	0.24
4	0.31	0.255	0.255	0.28	0.28	0.29	0.265	0.23

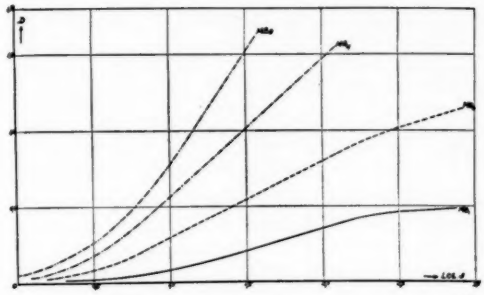


FIG. 16. Influence of the time of development on the density curve.

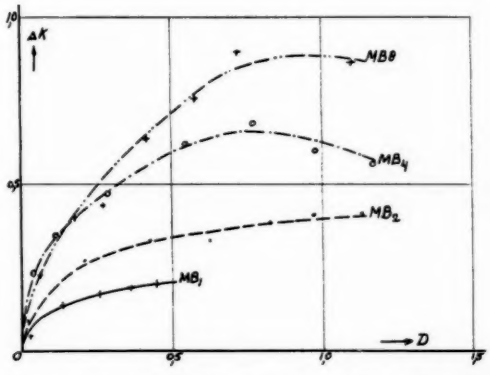


FIG. 17. Influence of the time of development on the absolute graininess.

Rapid plate. The figures are suppressed therefore.

2. For the metol developer (formula *A* of Table IV) we have studied the influence of the time of development. We used four Ilford Special Rapid plates, which were developed, respectively, for 1, 2, 4 and 8 minutes. The results are given in the Figs. 16–18. We see that ΔK increases with increasing time of development; this might be expected. The magnitude $\Delta K/\gamma_d$, however, is not influenced considerably. So we may conclude

TABLE VI. *Values of* $\Delta K_{max}/\gamma_t$.

	DEVELOPER							
PLATE	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*
1	0.85	1.57	0.90	0.68	0.88	1.39	0.77	0.49
2	0.60	0.60	0.51	0.37	0.53	0.55	0.44	0.37
3	0.48	0.51	0.53	0.41	0.59	0.40	0.42	0.27
4	0.39	0.40	0.37	0.34	0.35	0.37	0.34	0.28

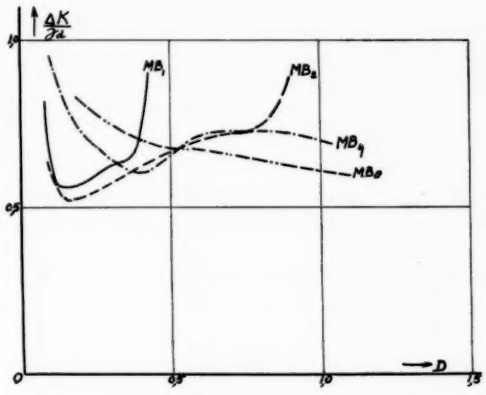


FIG. 18. Influence of the time of development on $\Delta K/\gamma_d$.

(apart from the properties of printing materials) that *the longest time of development is most favorable*, for it provides the highest speed and latitude, with equal graininess. It is possible that this conclusion will also hold good for other developers.

3. Finally we have studied some of the most recommended fine-grain developers. We used for this purpose two emulsions: Ilford Special Rapid (coarse grain) and Ilford Empress (finer grain), together with the three following developers:

Formula K. Agfa-"Atomal"-developer, according to the directions of the manufacturer.

Formula L. Physical developer of Odell.[8]

Formula M. Same as L, but with forebath of potassium-iodide, according to Odell.[8]

In general the results obtained with these developers are not superior to the formulae *A*, *B*, *C*, *D*, in some cases they are even inferior. We could not find a conclusive explanation for this discrepancy between the appraising statements of several authors and our measurements. We believe that in many cases the decrease of the contrast caused by the fine-grain developer is not taken in account; in other words that only ΔK, not $\Delta K/\gamma$ is considered. It will be necessary, however, to study more systematically the conditions of fine grain development, especially the influence of the particular emulsion used.

The authors are greatly indebted to Professor Dr. L. S. Ornstein for his interest in this paper.

[8] Cf. F. Odell, Brit. J. Phot. **80**, 303–305 (1933).

Capillary Chemistry and Fluorescent Light Microscopy*

E. A. HAUSER AND C. J. FROSCH†
Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received October 29, 1936)

THE fact that certain substances if irradiated with invisible light of short wave-lengths will emit visible light of characteristic colors and intensity (fluorescence or phosphorescence) has been recognized and used for a considerable time in analytical chemistry, crime detection, and similar suitable fields. The application of this phenomenon to microscopical research is of decidedly more recent date, because in microscopical work we have to look primarily for a light source giving a high surface luminosity, the total light efficiency being of secondary importance.

Although the mercury vapor lamp undoubtedly has a very high total light efficiency, the ultraviolet radiation depends solely on a few spectral lines which do not permit obtaining a sufficient surface luminosity to create a satisfactory fluorescence in extremely fine microscopic preparations.

A carbon arc, using carbons which have been impregnated with metal salts emitting a wide band of lines in the ultraviolet, therefore, had to be considered as a very important step forward. Yet since the greater part of the electric energy in such an arc is used up to maintain its temperature, comparatively little energy is available to emit the ultraviolet radiation. The light, therefore, is abundant with visible radiation and since the intensity of fluorescence is proportional to the amount of ultraviolet radiation absorbed by the preparation, substances exhibiting only faint fluorescence will not be detectable.[1, 3, 5, 6, 7, 10, 12, 13]

Recent constructions of so-called capillary (mercury) lamps, combining high total light efficiency and satisfactory surface luminosity might prove to be the final answer. For the time being, however, their fragility and comparatively very short life at full output, the need of special cooling jackets, etc., does not yet permit their general use.‡

The ideal light source for general use undoubtedly would be a pure metal arc, high in ultraviolet radiation. Of all the elements emitting ultraviolet light in an electric arc, iron seems to be the most satisfactory one from a common viewpoint, i.e., cheapness, simplicity of handling, avoidance of toxic fumes, etc. Using plain iron electrodes, nevertheless, has considerable disadvantages which are mainly the excessive formation of oxide and the danger of fusing the electrodes together. The difficulty of starting the arc must also be taken into consideration.

After a very careful and systematic study, M. Haitinger[4] has described a special arc which in conjunction with an appropriate optical system gives a maximum of ultraviolet radiation for microscopical work. The electrodes consist of an iron tubing filled with iron impregnated carbon. The filling permits an easy start of the arc and at the same time prevents excessive oxide formation on the electrode. Furthermore, it regulates the arc and permits maintenance of extreme steadiness. Both electrodes are placed in a vertical position so that the actual arc is located at right angles to the optical axis of the collecting and filtering system. Therefore, any excess of white light caused by the filter is enveloped by the pure iron spectrum of the tube type electrode so that the emitted radiation contains a maximum of the ultraviolet iron spectrum, which due to its broad band of lines results in very strong surface luminosity. To assure proper cooling of

* A paper delivered before the annual general meeting of the Optical Society of America on October 31, 1936, New York City.

† Present address: Bell Telephone Laboratories, New York City.

‡ *Note added at time of proofreading:* In the meantime we have been able to secure a General Electric type H3 capillary mercury vapor-lamp (air jacketed) as described by Saul Dushman in J. O. S. A. **27**, 17 (1937). A comparative spectrogram taken by D. Richardson at M. I. T. on a Bausch and Lomb medium-size quartz prism spectrograph of the Haitinger-Reichert iron arc, the capillary lamp and an ordinary low pressure mercury vapor lamp revealed the fact that the spectrum of the capillary lamp in the range between 3100 to 3900A emits a continuous spectrum with an extreme intensification of the dominant lines. This results not only in great light efficiency but also in excellent surface luminosity. Since this type of lamp, if run under normal conditions, has considerable life, it seems to be for the present the ideal light source for the work discussed.

the electrodes, which is essential for the maintenance of a quiet arc, they are placed in special brackets having a great number of cooling sheets. The arc housing is furthermore so constructed that satisfactory ventilation is guaranteed.

The light emitted from the arc or coming from the reflector passes through a collecting lens system and then through a cuvette containing a copper sulphate solution. The latter absorbs practically all the red light, thereby avoiding any excessive heat generation in the microscopic object. After passing the copper sulphate, the concentrated beam of light is filtered through a "Sendlinger UGI" filter, cutting out any radiation below 3100 and above 3900A. It then is directed onto the microscopic specimen, either by incident light optics or after being reflected by a mirror and passed through an appropriate condenser with a central dark field diaphragm is transmitted onto the microscopic slide. To avoid any possibly scattered ultraviolet radiation to reach the observer's eye, a safety filter is placed on top of the eyepiece, which filters out any light in the range of 3100 to 3900A, so that only light in the visible range can actually be observed.

Our knowledge in the field of the chemistry and physics of surface reactions of which the most generally known fall under the term "adsorption" has up to the present largely been based on results obtained by indirect experiments and deductions arrived from such data. The construction of a highly efficient fluorescent light microscope has finally made it possible to study visually some of these vitally important phenomena.

The problem of emulsification, for example, and the role which so-called emulsifying agents play in stabilizing emulsions has been unsatisfactorily explained, insofar as no methods were available to actually prove the theoretical assumptions put forward. With the use of fluorescent light microscopy, for example, we can immediately determine by simply stirring two liquids together which of the liquids has been emulsified, if at least one of the two exhibits fluorescence, or can be induced to fluoresce by dissolving in it substances of known fluorescence that are not miscible in the other phase. Such substances are generally known today as fluorochromes.[8] By adding emulsifying agents or substances that cause the formation of such emulsifiers in situ, we are now in a position to study the reactions actually taking place visually.[2] Other fields of interest are the problems of rate and degree of impregnation most important in a variety of industries. The entire field of ore flotation becomes of renewed interest as we now are in a position to actually follow the surface changes of the different phases of the system.

Wherever quantitative measurements are essential, the use of appropriate measuring instruments is commendable. So far, two methods seem to be available for this purpose. One can either use the ordinary photometer principle by applying a photometer eyepiece and photometrically determining any changes in intensity of the specimen to be investigated.[9] Or one can use an eyepiece containing a movable slit and connecting the eyepiece with a photocell recording equipment. The latter instrument, although possibly giving more accurate results, is more complicated as it naturally calls for a light source of constant intensity which, as explained in the beginning, is difficult to maintain. The use of the former instrument, for example, has already shown interesting results in determining the changes of concentration of capillary active substances in the interior of a liquid and in the neighborhood of its surface—in other words, it has been possible to visually and quantitatively study the accuracy of the Gibbs-Thompson rule with different systems.[11]

Literature References

1. E. M. Chamot and C. W. Mason, *Handbook of Chemical Microscopy* (Chapman and Hall Ltd., London, 1930).
2. C. J. Frosch and E. A. Hauser, Ind. Eng. Chem. Anal. Ed. **8**, 423 (1936).
3. M. Haitinger, Mikrochemie **9**, 430 (1930).
4. M. Haitinger, Mikrochemie **9**, 220 (1931).
5. M. Haitinger, *Die Anwendung der Fluoreszenzanalyse in der Mikrochemie* (E. Haim, Vienna (in print)).
6. M. Haitinger, *Fluoreszenzmikroskopie, Akad. Verlags. m.b.H.* (Leipzig (in preparation)).
7. M. Haitinger and H. Hamperl, Zeit. mikr.-anat. Forschg. **33**, H. 2, 193 (1933).
8. M. Haitinger and L. Linsbauer, Beih. Bot. Centralbl. **50**, Abt. 1, H. 2, 432 (1933); **53**, Abt. A. 378, 387 (1935).
9. E. Haschek and M. Haitinger, *Farbmessungen* (Emil Haim and Co., Vienna 1936).
10. A. Koehler, Zeits. wiss. Mikrosk. **21**, 129, 273 (1904).
11. J. W. McBain, T. F. Ford and D. A. Wilson, "New Methods of Studying the Surfaces of Ordinary Solutions," Kolloid Zeits. (in print).
12. J. A. Radley and J. Grant, *Fluorescent Analysis in Ultra-Violet Light* (Chapman and Hall, Ltd., London, 1933).
13. R. W. Wood, J. de phys., March (1919).

Photometric Discrimination with a Diffuse Boundary *

W. E. KNOWLES MIDDLETON
Department of Transport, Meteorological Service of Canada, Toronto
(Received January 11, 1937)

1. PURPOSE OF THE INVESTIGATION

THERE have been many determinations of the threshold of brightness-contrast under varying conditions of brightness,[1] angular extent,[2] and color[3] of field, all on the assumption of a fine, almost invisible line of division between the two parts of the field. The effect of a dark line of various widths upon the accuracy of photometric measurements has also been investigated.[4] But the writer has been unable to find any reference to a determination of the effect of diffusing the line of demarcation; and this is scarcely surprising, in view of the fact that no ordinary photometer provides a field of this sort.

A knowledge of this effect is therefore of no practical importance in photometry. In meteorology, however, it has an application in the theory of the visual range in fog.

In clean air or in haze, the visual range is a function of the contrast threshold ϵ, the extinction coefficient of the atmosphere, and the albedo of the object, which is supposed seen against the sky. In fog, there is a further effect due to the diffusion of light into the edge of the dark object. Fig. 1 will make the matter clear. The full line shows the distribution of brightness across the edge of a dark object seen at a distance in clear air; the dashed line shows the distribution in fog, the total contrast being supposed the same in both instances. In haze the transition will not be absolutely discontinuous, but will in general take place within an angle too small for the eye to appreciate. Only in a real fog composed of visible water droplets will the effect be marked.

It occurred to the writer that the effect of this diffusion, which has been recognized by Bennett[5] and taken into account by a separate term in an equation for the visual range, might be incorporated into the contrast threshold ϵ, this probably being a function of the angular extent of the diffusion. The aim of the present investigation is to determine this function.

2. THE BOUNDARY OF A DARK OBJECT IN A FOG

Suppose the boundary between a dark object and the sky to be represented by the line $x=0$ (Fig. 2), and consider this line as of infinite extent. From a narrow strip of the sky, dx in width, at $x=-\xi$, there will be diffused to $x=x_1$ an amount of light sufficient to produce a brightness

$$(dH)_{x_1}=H_h\cdot f(x_1+\xi)\cdot d\xi, \qquad (1)$$

where H_h is the sky brightness, supposed uniform. For the function f it is natural to assume the diffusion function

$$f(u)=ce^{-a^2u^2},$$

as has been done by Löhle;[6] and if we write

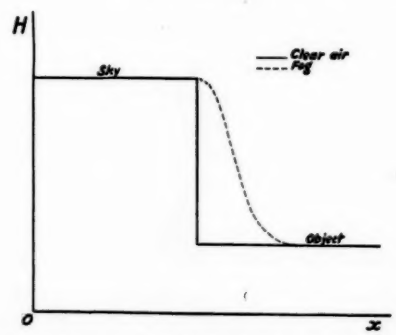


FIG. 1. Brightness distribution in clear air and in fog.

* Published by permission of the Controller of Meteorological Services, Department of Transport, Canada.

[1] E. M. Lowry, "Photometric Sensibility of the Eye and the Precision of Photometric Observations," J. O. S. A. **21**, 132–136 (1931).
P. G. Nutting, "Effects of Brightness and Contrast in Vision," Trans. I. E. S. **11**, 939–946 (1916).

[2] Helmholtz, *Physiologische Optik*, vol. 2, p. 286 (quoted by Koschmieder).

[3] A. Duschek, "Helligkeitsunterschiedsempfindlichkeit im Farbraum," Akad. Wiss. Wien., Berichte Abt. 2a **131**, 171–197 (1922).

[4] Y. Le Grand, "Sur la précision en photométrie visuelle," Rev. d'Opt. **12**, 145–159 (1933).

[5] M. G. Bennett, "The Physical Conditions Controlling Visibility Through the Atmosphere," Q. J. Roy. Met. Soc. **56**, 1–29 (1930).

[6] F. Löhle, "Lichtzerstreuung und Sicht," Met. Zeits. **46**, 49–59 (1929). It has been suggested by a referee that for the purposes of the experiments described below,

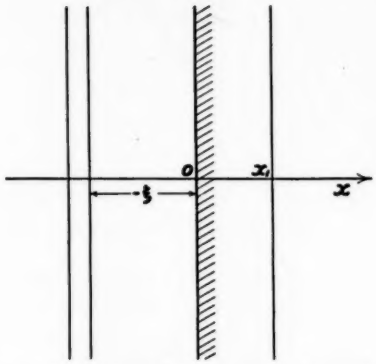


FIG. 2. To illustrate diffusion across a boundary.

$$\eta = a(x+\xi), \quad d\eta = a \cdot d\xi$$

and integrate from the boundary to minus infinity, we obtain for the total brightness at x produced by diffusion

$$H_x = H_h \cdot (c/a) \int_{-\infty}^{ax} e^{-\eta^2} d\eta. \tag{2}$$

We may express this as the difference of two integrals:

$$H_x = H_h(c/a)\left[\int_0^{\infty} e^{-\eta^2} d\eta - \int_0^{ax} e^{-\eta^2} d\eta\right] \tag{3}$$

or

$$H_x = (\sqrt{\pi} \cdot H_h \cdot c)/2a\left[1 - (2/\sqrt{\pi}) \int_0^{ax} e^{-\eta^2} d\eta\right], \tag{4}$$

in which the quantity within the brackets is the well-known error function, tables of which are everywhere available.

Furthermore when $x=0$, $H_x = H_h$ (the sky brightness), so that $c = 2a/\sqrt{\pi}$, and finally

$$H_x = H_h\left[1 - (2/\sqrt{\pi}) \int_0^{ax} e^{-\eta^2} d\eta\right]. \tag{5}$$

The parameter a, which appears in the limits of the diffusion integral, is thus a measure of the angular extent of the diffusion.

almost any reasonable transition would suffice. This is very probably true, and the discussion given in this section may be considered as of meteorological, rather than optical, interest.

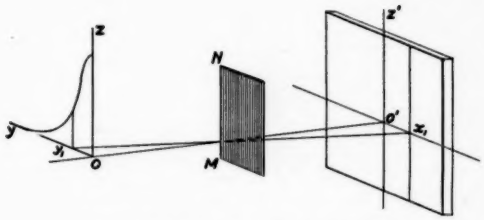


FIG. 3. Production of diffuse boundary.

3. THE APPARATUS

It was seen to be desirable to set up an apparatus to produce a bipartite field of which the diffuse boundary follows the law expressed in Eq. (5). Such a boundary could be produced by the shadow of a straight edge illuminated by a source of uniform brightness and suitable shape. Let such a source (Fig. 3) be bounded by the lines $y=0$, $z=0$ and a curve the form of which is to be determined. Let MN be the straight edge, and let it be parallel to OZ. Suppose a screen in a plane parallel to OY and OZ, and let the production of OM meet the screen in O'. Then a line element $x=x_1$ on the screen will receive light from all that part of the source lying farther from the z-axis than does the line $y=y_1$.

Let the curve bounding the source have the form $F(y)$. Then it can be shown very simply that the distribution of brightness given by Eq. (3) can be produced if

$$F(y) = K \cdot \frac{d}{dx}\left[\frac{2}{\sqrt{\pi}} \int_0^{ax} e^{-\eta^2} d\eta\right], \tag{6}$$

which can be evaluated by forming differences in a table of the error function.

This is strictly true only if the screen and the source lie in parallel planes, but is not seriously in error if the screen is rotated about $O'Z'$ as long as OO' is large compared to the dimensions of the source. The value of x corresponding to a given y is

$$x = (O'M/MO) \cdot y \cdot \sec\theta, \tag{7}$$

if θ is the angle between the two planes.

The actual apparatus is shown diagrammatically in Fig. 4. The screen X is set up at one end of an optical bench XL_2, and can be rotated about a vertical axis lying in its front surface and passing through the center line of the bench. On the

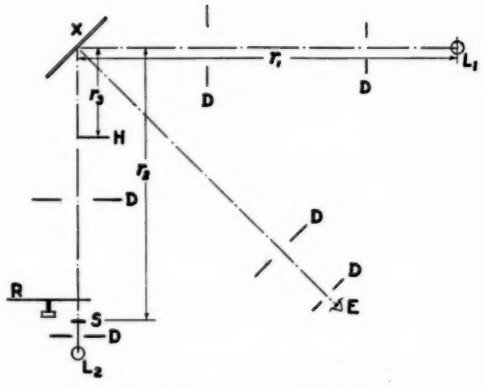


FIG. 4. Diagram of apparatus.

latter, the source S, consisting of a piece of finely ground glass placed in contact with the diaphragm described above is mounted 8.5 cm in front of a 250-watt spotlight lamp L_2, the assembly of lamp and source being capable of easy motion along the bench. H is the straight edge which casts the shadow; this is also mounted on a sliding carriage. R is a rotating sector for cutting down the illumination from S when desired.

The source L_1 illuminates X from a distance of one or two meters, and consists of a lamp similar to L_2. By using similar lamps run in parallel from the same supply, small fluctuations in voltage should be of no importance, as they should alter both the intensity and the color of the two sources equally.

The screen X consists of a board 30×30 cm, covered with black velvet, and having a removable central rectangle 7.5×12.0 cm which can be covered with any suitable material, such as white bristol board, and inserted so that its surface is flush with the rest of the screen. The stops D in front of L_1 are so proportioned that the whole of this white rectangle is uniformly illuminated by L_1. In order to orient X so that the light from the two sources should fall on it at equal angles, a small plane mirror was held against it with one edge bisecting the white rectangle vertically, and X turned until the corresponding edge of the beam from L_1 reflected by the mirror passed along the axis of the bench towards the edge of H. A viewing aperture was then set up at E along the normal to the screen and about 230 cm from it.

4. The Method of Experiment

It was first necessary to determine the ratio of the intensities of the two sources L_1 and S; this determination was carried out with the apparatus by a method similar to that of the Bunsen photometer, removing L_1 to a considerable distance (some 5 meters). The mean of 16 readings gave $I_s/I_1=0.00546$ with a probable error of about 1 percent.

The relation which it was desired to determine is between the two quantities, contrast threshold and width of diffuse boundary. Thus it would suffice to vary either the brightness ratio or the diffusion, keeping the other constant. With this apparatus, however, it is not possible to vary the brightness ratio in a continuous manner without the diffusion changing; consequently the brightness ratio was set at a series of known values and the diffusion varied by moving H.

As a measure of the diffusion it was decided to take the angle subtended at the eye of the observer by the distance between the line where the added brightness commences to fall off ($O'Z'$ in Fig. 3) and the line on which it has fallen to 1/100 of its initial value. The choice of this fraction is of course quite arbitrary. A simple calculation from the dimensions of the apparatus gives the following relation between the angle μ (in minutes) and the variable distances r_2 and r_3 (see Fig. 4).

$$\mu=24.4r_3/(r_2-r_3).$$

A simple nomogram was constructed for this relation.

By reversing H so that it covered the left half of the field, and moving it close to X, a sensibly sharp boundary, as in an ordinary photometric field, could be obtained. By varying the brightness ratio, the observer's threshold of contrast with no diffusion could be determined in this way.

The schedule for one complete experiment was as follows:

(1) Contrast threshold with sharp boundary—3 pairs of readings, alternately losing and picking up the boundary.

(2) Diffuse boundary—10 different brightness ratios, in systematic order, one pair of readings at a time, the whole series repeated three times, the second time in the reverse order of brightness ratios.

(3) Repetition of (1) in order to eliminate as far as possible the effect of fatigue, practice, etc.

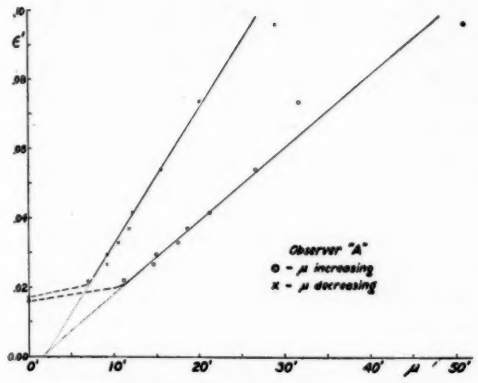


FIG. 5. Observer "A."

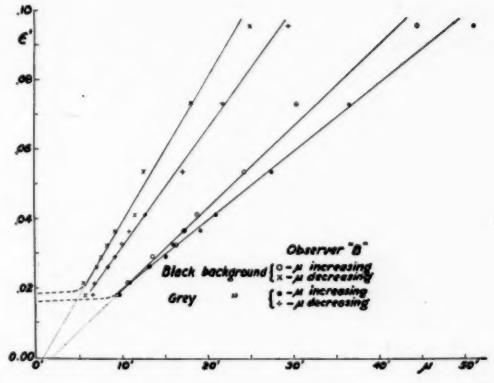


FIG. 6. Observer "B."

The observer signaled the appearance or disappearance of the boundary by means of a sharp blow on a hard object. The experimenter, who had been moving H (or S) in a slow and uniform manner, immediately stopped the motion on hearing the signal; it is certain that the personal error here involved was only a small fraction of the probable error of a single setting. The observer could not see the movements of the experimenter or form any judgment of the positions of H or S.

5. RESULTS

In view of the practical purpose of the investigation only a limited number of experiments have so far been performed. Four observers were tested; of these one seemed unable to come to any decision about the visibility of the diffuse boundary; another had no such difficulty but could not repeat results; the remaining two proved excellent in every way. None of the four were experienced in photometry, but all were used to other kinds of physical measurements, and were in excellent health.

The experiments were divided into two groups, one with a black background surrounding the field, and one with a neutral grey background about one-fifth as bright as the field. The brightness of the field was about 10 millilamberts.

The results are presented graphically in Figs. 5 and 6. Each point on these charts represents a mean of three independent experiments conducted on different days. The abscissae represent the diffusion in minutes as defined above, and the ordinates the contrast threshold. The points representing the disappearance of the boundary when the diffusion was being increased are shown by circles; those representing its reappearance with decreasing diffusion by crosses.

In such an experiment there is a restricted range of contrast where results can be obtained. Obviously one cannot work too close to the observer's ordinary (sharp boundary) contrast threshold; and when the contrast rises above about 0.06 the points begin to scatter badly. At about this contrast the two parts of the field are probably perceived to be of different brightnesses without reference to the boundary between them. Hence the lowest contrast used was 0.0218 (0.0186 for observer "B"); and while the experiments were at first continued up to 0.1652, only values up to 0.0957 have been plotted.

Within these limits the various sets of points seem to be represented best by straight lines; and it is interesting to note that for the black background the slopes of these lines are almost exactly the same for the two observers. The dotted extensions in the figures show that these lines do not by any means pass through any possible value for the contrast threshold with zero diffusion. Unfortunately, however, it was not possible to determine experimentally the form of that part of the curves near the axis of ordinates, and the dashed curves are purely speculative.

From the standpoint of the theory of visual range, the main result seems to be that the contrast threshold does not rise much until the

diffusion is at least five minutes of arc, using the curve which represents the finding of a boundary at first invisible. This is also true for the case of the grey background.

6. Discussion

It is to be expected that the effect of a diffuse boundary between the two parts of a photometric field will be different from that of a black line, such as has been investigated by Le Grand.[4] This author found a marked drop in the precision of photometric matches (i.e., a rise in the contrast threshold) as soon as the width of the black line attained one minute of arc; this change continued rapidly up to about 5′, thereafter proceeding more slowly. The initial large change in such a case would be due to the transition from the process of judging the disappearance of a boundary to the much more difficult process of judging the equality of two brightnesses.

In the present experiments, the junction between the two parts of the field was never visible as a definite boundary. It would therefore seem reasonable to suppose that, as actually occurred, there would be little rise in the threshold until the zone of transition was wide enough to be recognizable as such, i.e., until the observer could recognize that there was a part of the field with a brightness intermediate between that of the two main parts.

The writer can offer no explanation of the surprising linearity of the relation in the region between about 10′ and 30′. From some attempted observations with a contrast of about 0.165, which were not plotted on the diagrams, it seemed probable that the curve would become concave towards the axis of abscissae if extended; and indeed, this can be predicted from the simple consideration that there must be some brightness ratio which can be appreciated even if the two parts of the field are completely removed from one another, or even presented alternately to the observer. This point, however, is probably of no practical importance.

The fact that the threshold seems to remain about 0.02 until the boundary is at least 5′ wide seems to render it unlikely that diffusion has any great effect on the visual range of objects in fog, at least by way of altering the contrast threshold. The writer has been unable to find a record of any determinations of the width of the diffuse boundary produced by a fog; but it seems improbable that it could be wider than 5 minutes of arc. The differences[7] between the theory of Koschmieder and the experimental results of Bennett must therefore have some other explanation. Nevertheless some experimental work in natural fog, along the lines suggested, would be of interest.

It was the practical question just mentioned which prompted the writer to perform the experiments reported in this paper, and he realizes acutely their shortcomings, especially in the matter of the fewness of the observations. It was felt, however, that to pursue the matter to its conclusion was rather outside the proper sphere of a meteorological institute, when it could be done better with better apparatus in some laboratory of physiological optics.

In conclusion the writer wishes to acknowledge the kindness of several of his colleagues at the Meteorological Office in acting as subjects for his experiments, and of Professor G. P. Cosgrave of the University of Toronto in examining the arrangements from the standpoint of an experimental psychologist.

[7] M. G. Bennett, "Further Considerations Concerning Visibility by Day and Night," Q. J. Roy. Met. Soc. **61**, 179–188 (1935).

A Photoelectric Method for the Determination of the Parameters of Elliptically Polarized Light

C. V. KENT AND JAMES LAWSON
Blake Hall, Kansas University, Lawrence, Kansas

(Received January 2, 1937)

The given beam of elliptically polarized light traverses a phase compensator plate and *rotating* analyzer and enters a photoelectric cell suitably connected to an amplifier circuit and detecting device. The azimuth γ, of the polarizer producing the elliptically polarized light, and ϕ, that of the planes of polarization of the compensator (or the azimuth ϕ, and α, the phase difference introduced by the compensator) are adjusted until the detector (headphones) shows no modulation of the cell current. Tan γ, the ratio of the components of the elliptically polarized light, and Δ, the phase difference between them, can be calculated from the formulae; $\cos 2\gamma = \pm\sin\alpha \sin 2\phi$ and $\cot\Delta = \tan\alpha\cos 2\phi$. The method has been tested in the visible spectrum through a wide range of intensities. The sensitivity increases with the light flux in the cell and is practically independent of stray light. An accuracy of 1 part in 4000 or 5000 or less than 1 minute of arc in azimuthal readings is easily obtained. Polaroid plates can be used instead of Nicol prisms with no serious loss in sensitivity.

THE analysis of elliptically polarized light is usually accomplished with the aid of a phase compensator which plane polarizes the transmitted beam. The azimuth of an analyzer crossed with this beam, and the azimuth and phase difference of the compensator are then used to calculate the desired parameters. Reasonable accuracy can be secured only by the use of both azimuth and phase half-shade devices, but is severely limited at best by ever present stray light and eye fatigue at the low illuminations used.

These difficulties have suggested[1] the substitution of the photoelectric cell for the eye. Unfortunately, however, the characteristic virtues of the photoelectric cell are largely hidden in weak illuminations, although this disadvantage can be eliminated in part by adopting what is equivalent to a half-shade method. A quite different plan, which was contemplated by one of the writers for visual work many years ago, seems so peculiarly suited to the photoelectric cell that it has been developed and tested in this laboratory.

Briefly, in this method, the elliptically polarized light is reduced by the compensator *not* to plane polarized light, but to *circularly* polarized light which is passed through a rapidly rotating analyzer into a photoelectric cell. If, and only if, the beam is circularly polarized, will the photoelectric cell current be unmodulated by it, a condition which can be tested by well-known methods. The elliptical parameters are calculated from the azimuth and phase difference of the compensator. Thus, the photoelectric cell works at maximum illumination and stray-light effects become almost negligible. No half-shade devices are needed.

The equations used are closely analogous to those applicable in the visual method. The elliptically polarized light may be represented by its two perpendicularly polarized components, $\sin\gamma\cos x$ and $\cos\gamma\cos(x+\Delta)$ as in Fig. 1 (a), where γ is the azimuth angle, x the time angle ($2\pi ft$), and Δ the phase difference between the components. Passage through the compensator plate resolves each of these into components in the directions ϕ and $\pi/2+\phi$, and introduces a

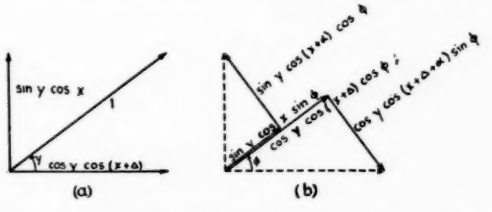


FIG. 1. Phase changes produced by the passage of plane polarized light through two phase plates. 1(a). 1, amplitude of original plane polarized light. γ, azimuth angle of planes of polarization of first phase plate. x, time angle. Δ, phase difference between components of light disturbance. Sin $\gamma\cos x$ and $\cos\gamma\cos(x+\Delta)$, components. 1(b). ϕ and α, azimuth angle and phase difference respectively of compensator phase plate. The arrows are the respective components of the light disturbance after passage through the plate.

[1] G. Bruhat and A. Guinier, Comptes rendus **12**, 396–416 (1933). G. Bruhat and P. Grivet, Compes rendus **199**, 718–720 (1934); **199**, 852–854 (1934).

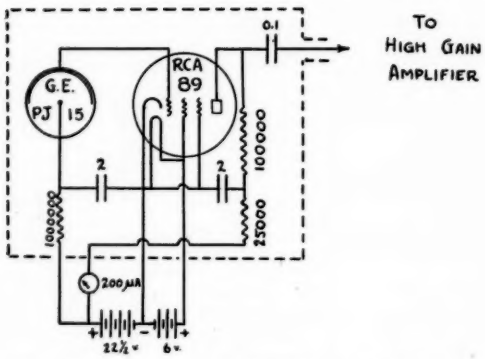


FIG. 2. Circuit of photoelectric cell and first amplifying tube. Resistances are in ohms, capacities in microfarads.

phase difference α between them in each case (Fig. 1 (b)). The resultant disturbances in these directions are then:

$$\begin{aligned}\Psi(\phi) &= \sin\gamma\cos x\sin\phi \\ &\quad +\cos\gamma\cos(x+\Delta)\cos\phi \\ &= (\sin\gamma\sin\phi+\cos\gamma\cos\Delta\cos\phi)\cos x \\ &\quad -\cos\gamma\sin\Delta\cos\phi\sin x \\ &= A\cos x+B\sin x.\end{aligned} \tag{1}$$

$$\begin{aligned}\Psi(\pi/2+\phi) &= \sin\gamma\cos(x+\alpha)\cos\phi \\ &\quad -\cos\gamma\cos(x+\Delta+\alpha)\sin\phi \\ &= (\sin\gamma\cos\alpha\cos\phi \\ &\quad -\cos\gamma\cos(\Delta+\alpha)\sin\phi)\cos x \\ &\quad +(\cos\gamma\sin(\Delta+\alpha)\sin\phi \\ &\quad -\sin\gamma\sin\alpha\cos\phi)\sin x \\ &= C\cos x+D\sin x.\end{aligned} \tag{2}$$

These are components of circularly polarized light if their amplitudes are equal and the phase difference between them is $\pi/2$. That is, if

$$A=\pm D, \quad B=\mp C. \tag{3}$$

Substituting values of A, B, C and D from (1) and (2) in (3) and eliminating Δ and γ from these two equations successively, we obtain

$$\cos 2\gamma = \pm\sin\alpha\sin 2\phi, \tag{4}$$

$$\cot\Delta = \tan\alpha\cos 2\phi, \tag{5}$$

and from (4) and (5)

$$\cos\Delta = \cot 2\gamma\cot 2\phi. \tag{6}$$

A knowledge of any two of γ, Δ, α and ϕ is sufficient to evaluate the other two. If the parameters γ and Δ of the elliptically polarized light are fixed, α and ϕ must be experimentally variable. Most experimental work will permit the variation of γ so that the compensator may be of the Stokes type with fixed phase difference and variable azimuth. Note that Δ and α are independently determinable so that the compensator constant α need not be known explicitly. The condition given by (5), $|\cot\Delta| < |\tan\alpha|$, can be met most conveniently by using a quarter wave plate in the compensator.

The optical set-up is identical with that of the simplest visual method except for the use of a rotating analyzer. This is mounted in a thin brass cylinder resting in ball bearings at its ends and rotated at a speed of about 40 r.p.s. by a belt over its middle from a small synchronous motor. The arm supporting the analyzer and motor is clamped to a stand entirely separate from the rest of the optical train so that no vibration will be transmitted to it. Small changes in position or direction of the axis of rotation of the analyzer have no effect on the readings.

The photoelectric cell and first amplifying tube are mounted in a metal shielding box which is set on a soft pad behind the rotating analyzer, its position not being critical. It is evident from Fig. 2 that fluctuations in the current common to the photoelectric cell and the grid of the first amplifying tube will produce corresponding fluctuations in the grid potential. These are subsequently amplified in a four-stage, high-gain amplifier, battery operated and designed to pass a modulation frequency of about 80. Sensitive headphones are convenient and satisfactory detectors. The simplicity of the circuit has permitted the reduction of ground noises to those due primarily to thermal agitation and the shot effect and it is doubtful if greater amplification would increase the signal to ground-noise ratio and therefore the sensitivity.

A most useful adjunct to the apparatus is a microammeter permanently connected in the plate circuit of the first amplifying tube and calibrated in terms of the light flux at the photoelectric cell. I/dI, where dI is the least observable light-flux modulation, has been found to increase linearly with the microammeter reading.

A Cooper-Hewitt "Lab-arc," a "Pointolite" and a 5-ampere carbon arc have been found to be satisfactory sources of light when operated

from storage batteries. Except with the Pointolite, commutator hum is prominent when these sources are operated from direct current mains. The green line of the mercury arc is isolated by a didymium glass filter, and in the case of the white light sources, a monochromator has been adjusted to pass a band of color 100 to 150A wide.

A fixed phase Stokes compensator has been used, primarily because of the ease with which mica and selenite plates, uniform over a large area, can be procured. The equations show that there are 4 symmetrical sets of polarizer and analyzer azimuths for which the no modulation condition holds. Differences between appropriate pairs of these give the 2γ and 2ϕ directly.

In testing the method, the light has been elliptically polarized by passage through a polarizer and any one of several phase plates, and the parameters measured for various colors and a 30-fold range of intensity. The technique of getting the γ and ϕ azimuths should be that of photometric balancing, i.e., quick, small passages of rapidly decreasing amplitudes through the no modulation position. The mean deviation of 10 successive readings has been taken as a fair measure of the sensitivity. This varies with the light flux from about 1 part in 600 to 1 part in 4000 or 5000 (less than 1 minute of arc) for the strongest illumination. The fraction is larger in the blue and violet owing to the high dispersion and absorption of the monochromator prism in this region.

In contrast to the visual method, stray light has a negligible effect, and weak daylight or direct current illumination, shielded from direct entrance to the photoelectric cell, is harmless. Due, no doubt, mainly to this independence, the azimuth settings for the 150A bands from the monochromator are as sharp as for the monochromatic light from the mercury arc. That is, in spite of the variation of the phase difference of the compensator plate with wave-length, the total effect of the band must be equivalent to that of a monochromatic circularly polarized beam plus a certain amount of unpolarized (stray) light. The parameters are average parameters and will vary if the relative intensities in the band are varied, say by changing the temperature of the source. This effect could be noticed.

The method is peculiarly adapted to the substitution of Polaroid or similar plates for Nicol prisms. These plates are not suited for accurate visual work because of the considerable portion of unpolarized light which is always transmitted; also, absorption, percentage of unpolarized light and the plane of polarization may vary across the disk. None of these factors have an appreciable effect when a Polaroid plate is used as the rotating analyzer, since the illumination through each element of area will be constant if the incident light is circularly polarized. They do produce an effect, however, when a Polaroid plate is used as the polarizer. But even in this case, the no modulation setting gives the true phase difference, Δ, and a mean polarizer azimuth which is a definite parameter, independent of intensity, stray light and wave-length. In agreement with these considerations, measurements of mean deviations with these plates are practically equal in accuracy to those with Nicol prisms for the same light flux at the photoelectric cell, reaching to better than 1 part in 4000.

It can be seen from the foregoing tests and discussion that the method is comparable with any yet devised in accuracy, except at low illuminations. It is superior in the simplicity of the optical train, absence of eye fatigue and stray light effects, and in being adaptable to the use of Polaroid plates. It is obvious that the method can be used to determine the azimuth of plane polarized light with the same high accuracy and convenience. In fact, routine measurements of this kind could be made automatic by the use of well-known relay circuits.

On the Precision of Photometric Observations

ALFRED H. HOLWAY
Biological Laboratories, Harvard University, Cambridge, Massachusetts
(Received January 7, 1937)

Lowry's results (1931) for brightness discrimination have been confirmed, his conclusion extended, and his recommendation amplified. The results of the present experiment, for example, show clearly that the root-mean-square departure $\sigma_{\Delta I}$ is directly proportional to ΔI. (This is here called Crozier's law.) Moreover, since ΔI is known to vary inversely with the size of the retinal field, the use of a test field which produces a large *retinal* area (central fixation) is to be recommended for increased precision in photometric work.

EVER since König and Brodhun's[5] classical study of brightness discrimination, workers in the fields of physical optics, sensory physiology, and psychology have exhibited an increasing interest in the form and behavior of functions defining the brightness sensibility of human subjects. This interest has reached a new peak in the recent, systematically exacting investigations of Hecht[3] and his associates. Until the work of Lowry,[6] however, little had been suggested in regard to the relationship now known to exist between fineness of discrimination and discriminatory precision for the human "eye," a relation that is of importance to the photometrician as well as to the experimental psychologist. The work reported in this paper was undertaken in an effort (1) to amplify Lowry's findings, and (2) to emphasize the nature and significance of the relationship which exists between discriminatory precision and fineness of discrimination.

The apparatus employed in the present investigation comprised a two-meter optical bench (calibrated in millimeters); appropriate screens; "neutral" filters (used in the observation telescope); a photometer head; two 500-watt, concentrated filament, Mazda projection lamps; a potentiometer control for the lamps; and an observer's cubicle. The photometric cube was of the equality-of-brightness (Lummer-Brodhun) type. By means of a Krüss reflecting prism, the light from the cube was redirected so that the eyepiece could be placed along the axis of the photometer-box. The total field presented to the eye was elliptical in shape. The major axis subtended, vertically, a visual angle of about 6 degrees; the inner field, also elliptical, subtended vertically an angle slightly less than 3 degrees at the eye of the observer. The exit pupil of the observation telescope was 2 millimeters in diameter; it thus constituted an *effective* artificial pupil.

The results of the investigation may be considered in two parts: Series I and Series II. In Series I, the observers were instructed to signify when the inner and outer fields were just noticeably different in brightness (method of "limits"); in Series II, to report when the two fields were equal in brightness (method of "average error"). In each series, the field of view was centrally fixated. Three practiced observers were employed throughout the experiment.

The photometer head and one of the lamp holders were secured in predetermined positions. The lamps had the same ratings in lumens per watt. The eye of the observer was always adapted for 6 minutes to the particular brightness level used. The procedure adopted in Series I was, starting at equality, first to increase the brightness of the inner field and to record the position of the lamp. The lamp was returned to the position of equality, and the inner field was then decreased. This general procedure was carried out in two ways: (1) by inverting the photometer head and (2) by reversing the direction in which the variable lamp was moved. Procedures (1) and (2) were used alternately throughout Series I. In Series II, starting alternately with the inner field either more or less bright than the outer field, the two fields were photometrically matched.

The same number of observations was secured at each brightness level in the two series. For a

specified procedure, the data are therefore homogeneous.[2] All brightness values were corrected for losses suffered in the photometer head and were finally expressed in terms of *retinal* illumination, i.e., in photons. Readings were thus obtained at definite intervals over a range extending from 1.00×10^{-3} to approximately 2.40×10^{3} photons.

TABLE I. *Results secured in Series I and II. For Series I, I_1 = brightness level in units of retinal illumination; $+\Delta I$ = just noticeable increment; $-\Delta I$ = just noticeable decrement in brightness. For Series II, $+I_E$ = photometric match when comparison was greater than standard during presentation; $-I_E$ = photometric match when comparison was less than standard. σ = root-mean-square variation. Each value is based on ten measurements.*

SERIES I: METHOD OF LIMITS
OBSERVER NO. 3

I_1 IN PHOTONS	$+\Delta I \pm\sigma$	$-\Delta I \pm\sigma$
2383.92	52.01 ±11.01	63.21 ±7.08
1101.14	29.72 ± 5.768	24.14 ±3.951
550.57	16.61 ± 1.744	15.20 ±2.002
275.29	12.08 ± 1.948	11.26 ±1.670
110.11	4.897 ± .5471	4.196 ± .872
55.06	2.951 ± .3174	2.515 ± .263
27.53	1.80 ± .2508	1.541 ± .3321
11.011	.871 ± .1790	.7112± .0915
5.506	.482 ± .0513	.4270± .0661
2.753	.315 ± .0641	.2575± .0322
1.101	.145 ± .0322	.1240± .0188
.5506	.1035 ± .0126	.1050± .0237
.2753	.07699 ± .0144	.0874± .0209
.1101	.053936± .0126	.0576± .00795
.05506	.03858 ± .00841	.0303± .00417
.02753	.02519 ± .00247	.0142± .00322
.01101	.01708 ± .00251	
.005506	.012008± .00152	
.002753	.007942± .00163	

SERIES II: METHOD OF "AVERAGE ERROR"
OBSERVER NO. 3

I_1 IN PHOTONS	$+I_E \pm\sigma$	$-I_E \pm\sigma$
2383.92	2441.44 ±18.21	2317.83 ±30.12
1101.14	1130.67 ±10.95	1074.81 ±15.75
550.57	565.39 ± 4.275	539.07 ± 3.621
275.29	281.602 ± 1.590	268.979 ± 2.505
110.11	113.521 ± 1.129	106.394 ± 1.416
55.06	56.645 ± .447	53.283 ± 1.000
27.53	28.291 ± .194	26.889 ± .251
11.011	11.3226 ± .0971	10.697 ± .0792
5.506	5.672 ± .0336	5.357 ± .101
2.753	2.833 ± .0231	2.682 ± .0399
1.101	1.1357 ± .00725	1.0796 ± .00710
.5506	.5723 ± .00487	.5260 ± .00731
.2753	.2879 ± .00451	.2445 ± .00418
.1101	.11875± .00210	.0974 ± .00614
.05506	.06263± .00135	.05145± .00108
.02753	.03242± .00174	.02477± .000792
.01101	.01321± .000690	.00918± .000863

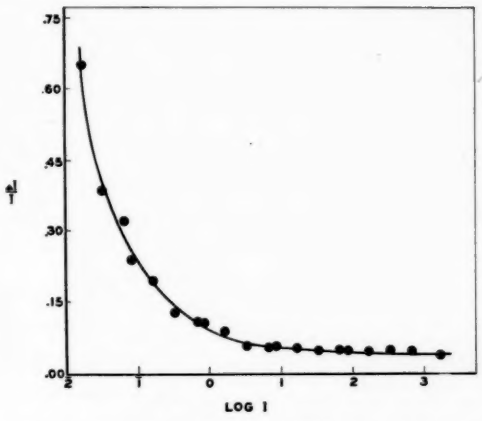


FIG. 1. The Weber fraction, $\Delta I/I_1$, as a function of log I_1 for Observer No. 1 (data from Series I): $\Delta I = I_2 - I_1$ is the just noticeable increase in retinal illumination expressed in photons; I_1 is the level of illumination employed. The Weber ratio, $\Delta I/I_1$, decreases rapidly at first and then tends to approach a minimal value. The functions for the other observers are essentially identical in form.

RESULTS

The data obtained in Series I and II for Observer No. 3 are presented in Table I. (Those for Observers No. 1 and No. 2 will be published elsewhere.) Each value is based on ten measurements; all values are expressed in units of retinal illumination (photons). The measurements for Series II depict the relation between level of illumination and the variation in photometric matchings. The values for Series I have to do with least discriminable differences, and are basic to the subsequent discussion.

Results secured in Series I are also presented graphically. Fig. 1 shows graphically the relation existing between $\Delta I/I_1$ and log I_1 for Observer No. 1. The curve definitely approaches an asymptote at high intensities. The functions for the other observers are of the same general shape. Lowry's function, on the other hand, clearly exhibited a tendency to rise at high intensities; the lower limb of the function shown in Fig. 1 does not manifest this tendency. In either set of experiments, it is likely that the observers were sufficiently well adapted at high intensities. The discrepancy is probably not due to the factor of adaptation.*

* The difference in the units employed raises no difficulty. Our units (photons) may be converted into millilamberts by the simple process of dividing our abscissa values by a

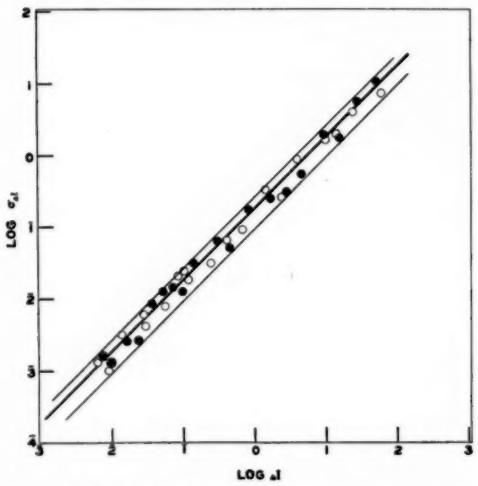


FIG. 2. Mean ΔI is directly proportional to the root-mean-square departure (Crozier's law). The central line is drawn so that it bisects arithmetically the vertical distance between the enclosing limits. The line has a slope = 1. Open circles, ΔI for increment in I_1; solid circles, ΔI for decrement in I_1. The data are for Observer No. 3. Cf. also, Fig. 3.

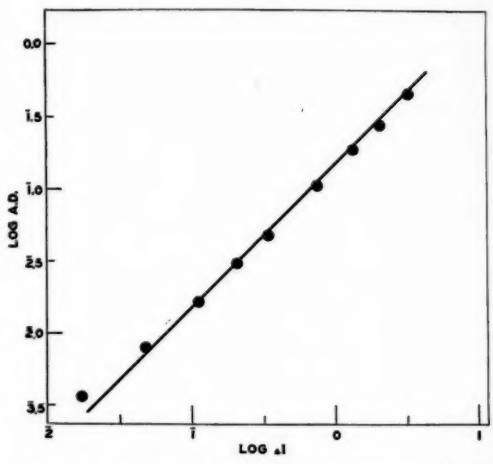


FIG. 3. Lowry's data (1931) for six observers showing correlation between ΔI and average deviation at fixed values of I_1. The double-log grid is extremely sensitive to small variations at very low values. The line drawn in this figure has a slope = 1. See text.

There remain at least two other possibilities: (1) Lowry's maximal intensities differed slightly from ours, and might have produced other than visual effects (reference 5, p. 918); (2) the presence or absence of what has been called the *surround* might have been responsible (the cube in Lowry's photometer possessed a vertical dividing line, whereas the one used in the present work had concentric fields). The first possibility could be entertained on the basis of the effects reported by König and Brodhun in 1888. However, since Lowry did not report the occurrence of such disturbing effects, and in the light of the more recent researches of Hecht (reference 3, p. 763), this possibility may be rejected. The second possibility quite probably provides the *raison d'etre* for the discrepancy. It is to be expected that the use of a surround (even though, as in our case, a small one) would tend to "level off" the curve at high intensities.† Accordingly, both sets of results can be brought into harmony with laboratory expectation.

On the other hand, since the $\Delta I/I_1$ *vs.* log I_1 functions obtained in the two laboratories do differ (Lowry's function, like that of König and Brodhun, tends to rise at high intensities; ours does not), and since the coefficients of dispersion for both functions are known, the two sets of results provide an immediate test for what may be appropriately called Crozier's law.[1] In general form:

$$d(\sigma_{\Delta I})/d(\Delta I) = K \cdot (\sigma_{\Delta I}/\Delta I), \qquad \text{(A)}$$

where $\sigma_{\Delta I}$ is the root-mean-square deviation of ΔI, the just discriminable interval; K is a constant. When K is found to be equal to unity, then $\sigma_{\Delta I}$ must be directly proportional to ΔI: i.e.,

$$\sigma_{\Delta I} = k'\Delta I. \qquad \text{(B)}$$

This law (which is in reality an experimental finding) provides an effective means for the determination of certain properties of the operationally defined variables $\sigma_{\Delta I}$ and ΔI.[1, 2, 8] An analysis of Fig. 2, for example, indicates that ΔI may not be the independent, but the dependent

factor of 10. This follows from the fact that our artificial pupil has a *radius* equal to unity. The conversion of Lowry's units into photons, on the other hand, can be approximated by making a pupillary "correction." When this correction is made, it is found that Lowry's brightness levels range from 1.60 up to 3.65 (expressed in log photons).

† These comments are, of course, in no sense to be construed as statements subversive of the experimental excellence manifested in Lowry's investigation. Photometers without "surrounds" inevitably provide a region (ideally a point) at which the Weber fraction, $\Delta I/I$, is minimal (cf. König and Brodhun). Lowry's results are exactly what one would predict; they constitute, therefore, a direct confirmation of expectation.

variable in the above relation. (See Fig. 3 for a plot of Lowry's data.) The general form of this equation is known to be valid for the human observer in the *kinesthetic*, *tactual*, *auditory*, and (now) the *visual* discrimination of intensities.[2a] The implications of the equations are therefore of general as well as particular theoretical interest.

Quite apart from theory, however, Eq. (B) has implications of a very practical nature for the photometrician. Since ΔI and field size are related by a function that is hyperbolic in form,[4, 7] it appears that the larger the illuminated *retinal* area, the greater the precision of observation for any specified brightness level. This follows directly from Eq. (B).

It is true, of course, that Eq. (B) has to do with "absolute" values, and illuminating engineers are ordinarily interested in the percent that the variation is of the mean, i.e., in reducing the *relative* variation to a minimum. As noted above, Lowry has defined the region of minimal relative variation for a particular cube and specified field size. It follows from further consideration of Eq. (A) and Fig. 1 that the relative variation should also decrease with an increase in the size of the retinal area employed.

It is a pleasure to acknowledge my indebtedness to Professor Lucy Wilson of Wellesley College for the use of laboratory space and several pieces of apparatus.

Bibliography

1. W. J. Crozier, J. Gen. Physiol. **19**, 503 (1935–36).
2. W. J. Crozier, Proc. Nat. Acad. Sci. **22**, 17 (1936).
2a. W. J. Crozier and A. H. Holway, Proc. Nat. Acad. Sci. **23**, 1 (1937).
3. S. Hecht, Hdbk. Gen. Exper. Psychol. 743 (1934).
4. S. Hecht, Cold Spring Harbor Symposia on Quantitative Biology, III (1935).
5. A. König and E. Brodhun, Berlin Akad. d. Wissen. Sitzsber. **37** (2), 917 (1888).
6. E. M. Lowry, J. O. S. A. **21**, 132 (1931).
7. J. Steinhardt, J. Gen. Physiol. **20**, 185 (1936).
8. M. Upton and W. J. Crozier, Proc. Nat. Acad. Sci. **22**, 412 (1936).

A Further Search for the Ideal Color System

I. A New Mechanico-Graphical Method

R. H. SINDEN
Wolcott, New York

(Received December 24, 1936)

I. INTRODUCTION

A NUMBER of attempts have been made to correlate on theoretical grounds data pertaining to chromatic thresholds[1] with the trichromatic coefficients of the spectral and nonspectral colors. The earliest was that of Helmholtz, who sought to establish a relationship between these two distinct sets of experimental facts on the basis of the so-called Young-Helmholtz theory of color vision. Starting with the assumption that the logarithmic law of sensation intensity (the Weber-Fechner principle) holds individually for each of three independent specific sensations of color, he developed an expression for the chromatic threshold as a function of relative differences of the elementary sensation intensities. Then he endeavored to show that the equation he had set up agreed with observation. To do so would be not alone to prove the theses upon which it was grounded, but also to provide a means of setting up a color system resolved upon the true ultimate variables of the color sense—an objective of practical as well as of theoretical consequence.

In comparing his equation with the results of experiment, Helmholtz was obliged to depend upon certain numerical data of König's, although these were not as accurate as could be desired and moreover did not cover the full range of the visible spectrum. Nevertheless, within the restricted spectral range, he was able to obtain results that seemed to show a measure of agreement. Supposing the discrepancies were largely ascribable to experimental error, he concluded that the proposed relationship or an analogous one exists.

In recent times, the assembly of accurate and comprehensive data pertaining to the intrinsic characteristics of the color sense, coupled with the growth in practical importance of the science of colorimetrics, has awakened a new interest in the problem propounded by Helmholtz. Before considering modern developments a brief resume of the classical treatise is desirable.

The general equation developed by Helmholtz[2] from considerations revolving about the Weber-Fechner principle took the form (in quadrature):

$$dE^2=\left(\frac{dx}{x}\right)^2+\left(\frac{dy}{y}\right)^2+\left(\frac{dz}{z}\right)^2, \quad \text{(H-1)}$$

in which dE denotes the magnitude of the effect upon the consciousness of simultaneously changing the excitations x, y, z of the three primary sensations (in the sense of the Young-Helmholtz theory) by small amounts dx, dy, dz, respectively.

When the difference in chromaticity between two colors nearly alike is measured under equality-of-brilliance conditions, this gave for the sensation of their difference:

$$dE^2=\frac{1}{3}\left[\left(\frac{dx}{x}-\frac{dy}{y}\right)^2+\left(\frac{dy}{y}-\frac{dz}{z}\right)^2+\left(\frac{dz}{z}-\frac{dx}{x}\right)^2\right]. \quad \text{(H-2)}$$

In order to put Eq. (H-2) into a convenient form for the substitution of numerical data, Helmholtz next introduced the factor $\delta\lambda$,[3] replacing dx, dy, dz by $dx/d\lambda$, $dy/d\lambda$, $dz/d\lambda$. He then showed that the x, y, z of the above expressions

[1] The writer has chosen to adopt in this discussion the nomenclature suggested and defined in the Report of the 1920–21 Committee on Colorimetry of the Optical Society of America. Accordingly, the term *brilliance* supplants the more cumbersome "apparent brightness," and its complement, *chromaticity*, designates the attributes of the generic sensation of color apart from brilliance. Thus, "chromaticity" has a wider significance than "hue," which is usually looked upon as complementary to "saturation." *Chromatic threshold*, therefore, seems to be the proper descriptive term for the quantity under consideration, of which the hue-difference and the saturation-difference thresholds are in a sense special cases. Cf. J. O. S. A. and R. S. I. **6**, 527 (1922).

[2] H. v. Helmholtz, Physiol. Optik, 2 ed. pp. 444–458.

[3] "The wave-length threshold of chromaticity-difference discrimination in the spectrum."

are connected with the spectral "trichromatic coefficients," commonly designated by R, G, B, by linear homogeneous equations of the type

$$\left.\begin{aligned} x &= a_1R+b_1G+c_1B \\ y &= a_2R+b_2G+c_2B \\ z &= a_3R+b_3G+c_3B \end{aligned}\right\} \qquad \text{(H-3)}$$

whose coefficients, however, were at the start unknown.

The problem then was to determine the coefficients of these equations of transformation in such a manner as to give equal values of dE for a series of spectral wave-lengths, when the corresponding numerical values from the data of König were substituted for R, G, B and $\delta\lambda$. Since the choice of units for x, y, z is arbitrary, only two parameters for each equation had to be definitely fixed.

The solution was accomplished by the method of successive approximations; i.e., different numerical values were substituted for a_1, b_1, c_1, etc. until, by trial and error, those values were found that made the quantities dE most nearly alike for all the wave-lengths considered. The ultimate values for the coefficients of Eqs. (H-3) thus determined were then used to convert König's spectral trichromatic coefficients to a new system in which the primaries x, y, z presumably corresponded to the elementary sensations of color. Finally, the x, y, z coordinates of the spectral colors were plotted in a trilinear-coordinate system and thus the spectral locus established with respect to the positions of the "fundamental sensations," which lay at the apices of the parent triangle.

To demonstrate the agreement of his results with observation, Helmholtz plotted calculated values of $\delta\lambda$ together with König's experimentally obtained values. The curves constructed from the two sets of figures show a close similarity in form (although the theoretical curve has been "smoothed" to such an extent that some of the computed points lie at considerable distances from it).

Comparatively recently, Schrödinger,[4] in the course of an exhaustive analytical treatment of the basic principles of colorimetrics, showed that Helmholtz's "dE" may be looked upon as a line element in a three-dimensional continuum of which x, y, z are rectangular coordinates. He then undertook to demonstrate that the Helmholtz "line element" (Eq. (H-1)) leads to a family of curved surfaces of uniform brilliance in the "color space;" whereas experiment proves that surfaces of uniform brilliance are planes. In order to circumvent this difficulty Schrödinger set up a modified line element analogous in form to the Helmholtzian expression, which included the factor $1/h$ (the reciprocal of luminosity).

In testing this equation, he did not proceed, however, as Helmholtz had done, to determine the x, y, z so as to obtain the best possible coordination among the available experimental data, nor did he accept Helmholtz's "fundamental sensations," but instead, set up beforehand a set of primaries based on certain arbitrary considerations. Using Helmholtz's computed data modified to suit the primaries of his choice, and values for $\delta\lambda$ from Uhthoff, he computed dE for a limited number of points in the spectrum and showed that the mutual agreement of these quantities was about on a par with that obtained by Helmholtz.

Bouma and Heller[5] have since compared Schrödinger's line element with more recent observational data, extending the comparison farther toward the extremes of the spectrum and also investigating its conformity to saturation-sensibility figures. Finding the agreement unsatisfactory, they suggested still another alternative form. By separating the brilliance and chromaticity elements, they arrived at the generalization:

$$dE^2 = (dh/h)^2 + f(x, y, z, dx, dy, dz),$$

where the second term on the right is a particular kind of function such that when the proportions $dx:dy:dz = x:y:z$ hold, its value is zero. This being the case, in a transition from one color to another of the same excitation type (e.g., when the intensity of a stimulus of fixed spectral character is altered) the second term vanishes and the first gives the Weber-Fechner law. On the other hand, on a plane of uniform brilliance the first term disappears and the second gives the chromaticity differential. Since measurements of

[4] Ann. d. Physik **63**, 481–520 (1920).

[5] Proc. Roy. Acad. Sci. Amsterdam **38**, 258–271 (1935).

the chromatic threshold are always made under equality-of-brilliance conditions, thus in dealing with actual data the first term can be altogether neglected.

Bouma and Heller then tried certain specific forms of $f(x, y, z, dx, dy, dz)$, selecting arbitrary primaries and using in their computations the best modern data; but no marked improvement over the earlier results was gained, and the conclusion was reached that a satisfactory chromaticity-element of this type cannot be found.

In addition to these attempts to correlate colorimetric data on a theoretical background, a universal empirical formula for this purpose has been proposed by Judd,[6] of the form

$$KdE=dD \qquad (K \text{ is a positive constant}),$$

where D is distance in any direction on a trilinear color chart or "color triangle" constructed with certain specified primaries. This leads to a linear relation for the difference in chromaticity between colors differing in finite degree. To denote such differences in chromaticity, Judd introduces the symbol ΔE and writes

$$K\Delta E=D.$$

The constant K may be chosen so that ΔE is equal numerically to the number of just perceptible differences of chromaticity, or *unit chromaticity steps*, between two colors separated by a distance D on the given color triangle.

Judd has determined the coordinate system most advantageous for his expression and has shown that it is capable of giving results in close agreement with chromatic sensibility data of various kinds. For data obtained in separate investigations the constant K must in general assume different values, since the absolute magnitude of the chromatic threshold varies widely according to experimental conditions and methods, as well as to some extent among individual observers. The agreement with experiment is very much better than had previously been obtained. Judd estimates that about ninety percent of available data on chromatic sensibilities are represented within their experimental uncertainty. The extreme simplicity of the relationship is an asset in the field of applied colorimetrics, in which it is primarily intended to find application.

[6] J. O. S. A. **25**, 24 (1935).

It does not necessarily follow, however, that Judd's equation gives the closest possible representation of experimental fact, or that it is of the form most likely to lead to advancement in a theoretical direction. From the point of view of the theorist, the Helmholtz chromaticity differential in its general and special forms given by Eqs. (H-1) and (H-2), which alone is derived from fundamental concepts by a connected sequence of logical steps, remains in certain respects the most interesting that has yet been proposed. The objection brought against it by Schrödinger on theoretical grounds may turn out to be of little importance. For example, if it should be found that the relative amounts of the three elementary color sensations never differ more than a little even for the colors of the pure spectrum, then Schrödinger's criticism is of academic interest only. As a matter of fact there is reason to suppose in advance this may prove to be the case. Troland,[7] Hecht[8] and others among present day visual theorists have amassed considerable evidence indicating that the physiological saturations of the spectral colors are much lower on an absolute scale than was formerly held. Thus, from a theoretical standpoint no serious hindrance to a reconsideration of the Helmholtzian functions appears to exist.

The numerical results obtained by Helmholtz in his investigation of Eq. (H-2), it is true, can scarcely be regarded as other than negative if put to the test of a comparison with modern colorimetric data. However, in his search for the primaries conducive to optimum results with this expression, Helmholtz, as we have seen, was severely handicapped by lack of adequate experimental data. A new attempt employing modern data could hardly fail to better his result to a certain extent. Accordingly, this has been undertaken. The outcome has proved fortunate to an unanticipated degree, and shows that Helmholtz was led far afield by the deficient data he was constrained to use, failing even to approximate an ideal solution of the problem. A close parallelism between the consequences of Eq. (H-2) and experimental fact is revealed.

The arithmetical trial-and-error procedure followed by Helmholtz in the selection of his

[7] J. O. S. A. and R. S. I. **6**, 3 (1922).
[8] J. O. S. A. **21**, 615 (1931).

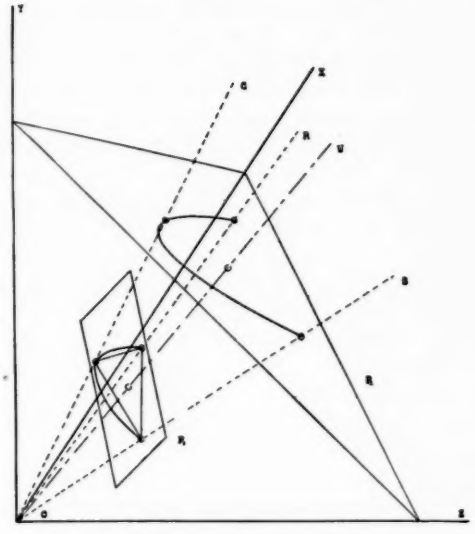


FIG. 1. Geometry of the color space.

primaries involved an immense amount of numerical computation, even though relatively few experimental points were taken into consideration. In order to obviate the necessity for tedious calculations and at the same time to permit a more efficient utilization of the available experimental data, a new method of approach has been devised. Although particularly adapted to the present problem, this mechanico-graphical method may with modifications be applicable to related problems in other fields. In the remaining portion of this paper, the method will be described, together with an apparatus which has been constructed to put it into effect, and the application of the latter to the solution of the special problem in hand. A later paper will present results already obtained with its use.

II. THEORY

The customary way of representing a color system is by a two-dimensional diagram, the "color triangle." While the color triangle is adequate for many purposes, a more comprehensive method of representation is by means of a color space, that is, a three-dimensional rectangular coordinate system in which the coordinates correspond to varying amounts of three primary colors.

At the outset let it be agreed that all color match equations shall hold throughout the entire color space; in other words, that the "Purkinje phenomenon," which affects only a restricted region close to the origin of coordinates, shall be neglected. In a system of this sort the locus of colors of like excitation type is a straight line emanating from the origin—the point of zero intensity—and extending on indefinitely. The direction of the line with respect to the coordinate axes defines chromaticity, while the distance from the origin defines the excitation intensity at any particular point on the line. In such a system the choice of "primary" colors to be assigned to the axes of coordinates is arbitrary, just as is the choice of primaries in the color triangle. For the purposes of the present discussion, however, let a color space be assumed in which the axes represent the natural primaries of the color sense mechanism, which, following Helmholtz, will be designated by x, y, z.

In this color space all real colors are contained in the positive octant because in the sense of the Young-Helmholtz theory x, y and z cannot have negative values. The spectral hues therefore are represented by a group of straight lines lying in the positive octant which taken together form a cone-shaped open surface with its apex at the origin. But the positions of these lines are not fixed *a priori* with respect to the coordinate axes, since they depend upon the choice of units in which x, y and z are stated. For example if the units are chosen so that for the standard white, $x=y=z$, then this white, W, will be represented by the median line of the octant, and the spectral surface will occupy a corresponding fixed position with reference to the axes.

The color triangle can be derived from this three-dimensional system. Thus the spectral primaries R, G, B, in terms of which the trichromatic coefficients of the spectral colors are known, are represented by three lines on the spectral surface. Let a plane, P_1, Fig. 1, be passed through the latter in a direction such that the points where these three lines pierce it lie at the apices of an equilateral triangle. Then the curved line traced on the plane by intersection with the spectral surface is the spectral locus in an equilateral-triangular coordinate system for which R, G and B are the primaries.

The white point of the diagram is the point at which the plane is pierced by the white line of the color space, and will not in general lie at the center of the R, G, B triangle. By a suitable choice of units for x, y, z, it may, however, be made to lie at the center or at any other point in the triangle. A corresponding change in the spectral locus would of course occur. When the units for x, y, z are chosen so that the white point lies at the center of the triangle, the diagram thus obtained must coincide with one constructed in equilateral-triangular coordinates using the standard spectral trichromatic coefficients, since, given the three primaries and the location of the white point, only one spectral locus is possible.

Now let a second plane, P_2, Fig. 1, be passed through the positive octant so as to intercept the coordinate axes at equal distances from the origin. The lines of intersection of this plane with the coordinate planes will form an equilateral triangle with apices at the piercing-points of the axes. The spectral surface will intersect the plane in a curved line, which must in this case lie wholly within the triangle. The diagram thus delineated is a color triangle with the natural variables x, y, z as primaries and with the spectral locus shown in its true relation to the latter.

The three lines on the spectral surface denoting the spectral primaries R, G, B will pierce the plane at three points on the spectral curve. With the locations of these three points known with reference to the x, y, z triangle, it then would be a simple matter to evaluate the set of constants a, b, c in Eqs. (H-3) and thence to find the trichromatic coefficients of all colors in terms of x, y, z. Under the circumstances the white point of the second diagram will not in general lie at the center of the x, y, z triangle. However, if the constants a, b, c are known for one set of units for x, y, z, they can readily be found for any other set, in particular for that set which will cause the white point to fall at the center of the x, y, z triangle.

III. Method

The foregoing considerations should help to make clear the working of the device next to be described. At the origin of a three-dimensional coordinate system is placed a point source of light. A translucent screen is mounted so as to intercept the coordinate axes at equal fixed distances from the origin. Between the screen and the light source is introduced a transparent glass plate mounted on a universal holder so that it can be made to occupy any desired position in the intervening space. On the glass plate is drawn in black ink the spectral locus obtained by plotting the standard spectral trichromatic coefficients in equilateral-triangular coordinates. The positions on the spectral curve of the three spectral primaries R, G, B are marked, likewise the location of the white point.

When the light is turned on, the curve on the glass plate casts its shadow on the translucent screen. By suitable manipulation of the glass plate it is possible (in the ideal case) to make the shadows of the three points R, G, B on the spectral locus coincide with any three given points on the screen. That is equivalent to assigning to the remaining six (three of them are arbitrary) of the nine constants a, b, c of Eqs. (H-3) any desired numerical values. Therefore, it is surely possible to find a position of the glass plate such that the shadow on the screen will represent the spectral locus in its true relation to the natural primaries x, y, z. The problem is to determine that position.

According to the hypothesis under test, the problem will have been solved if an adjustment can be found such that dE as given by Eq. (H-2) assumes the same value at every point on the shadow locus. This equation may be restated in somewhat simpler terms. The symbol dE as used by Helmholtz is in a sense ambiguous, in that it denotes at different times an infinitesimal variable and a finite fixed quantity. To avoid this confusion, the term δE will henceforth be adopted to designate the fixed sensation value of the chromatic threshold. Since the unit in which δE is stated may be arbitrarily chosen, an indeterminate factor is introduced into the equation, to which may be assigned any convenient value. The expression then becomes

$$K^2(\delta E)^2 = \left(d \log \frac{x}{y}\right)^2 + \left(d \log \frac{y}{z}\right)^2 + \left(d \log \frac{z}{x}\right)^2. \quad \text{(H-2a)}$$

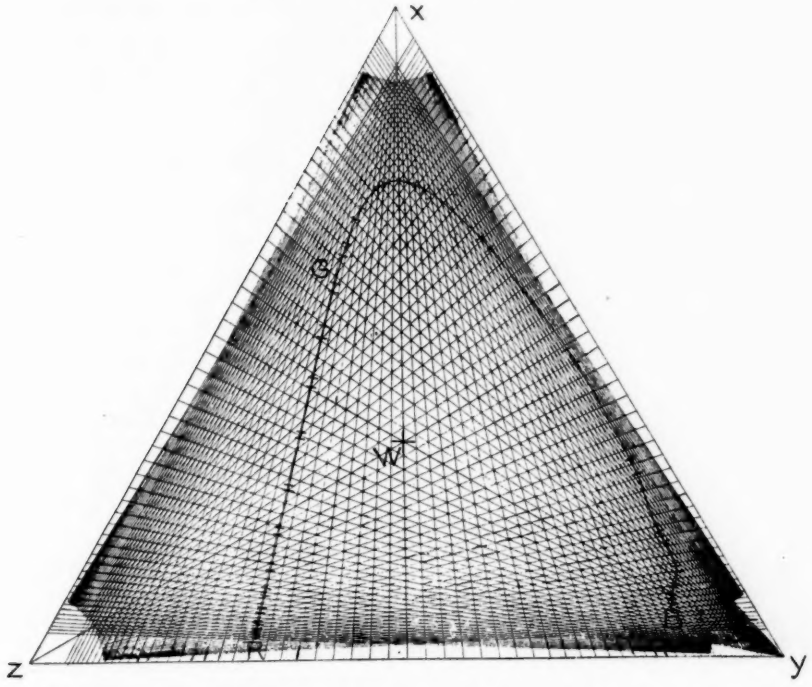


FIG. 2. The special coordinate system employed on the screen of the projection apparatus, in which the coordinates are log x/y, log y/z and log z/x. Coordinate interval, $\Delta \log_e = 0.1$. A spectral curve with $\delta\lambda$ segments is shown.

If the three logarithmic quantities are allowed to be represented respectively by u, v, w, the equation may be written

$$K^2(\delta E)^2 = (du)^2 + (dv)^2 + (dw)^2. \quad \text{(H-2b)}$$

In order to test whether the correct setting has been obtained it is necessary, then, to determine relative values for δE at various points along the shadow locus. To facilitate this, the curve on the glass plate has marked off on it at wave-length intervals of 10 mμ along its extent, segments, the lengths of which, in wave-lengths, are equal to experimental values of $\delta\lambda$ at the corresponding points. On the screen coordinate lines are drawn, so that the values of $\delta\lambda$ as indicated by the lengths of the segments in the projected curve may be translated into terms of x, y, z. The coordinates of the system are not x, y, z, however; instead, they are u, v, w (where $u = \log x/y$, etc.). Thus the differences of the coordinates of the two ends of a segment give directly the corresponding values of du, dv and dw. A simple computation then gives δE. In practice, moreover, as will presently become clear, an approximate idea of the relative values of δE may be had by mere inspection.

It will now be necessary to glance briefly at some of the characteristics of the special coordinate system employed on the screen of the projection apparatus. A system of this type is shown in Fig. 2. First, it should be noted that at every point on the diagram the sum of the three logarithmic coordinates u, v, w is zero. Now let any two points p_1 and p_2 have coordinates, respectively, u_1, v_1, w_1 and u_2, v_2, w_2. Granted that the value of the chromatic threshold is everywhere given by Eq. (H-2b), it is desired to find the number of just perceptible steps of difference in chromaticity between p_1 and p_2. Take first the case where both points lie on the same line of the logarithmic system, for instance on one of the u lines, so that $u_1 = u_2$. Then in traveling the shortest distance from one point to the other, for each unit difference in

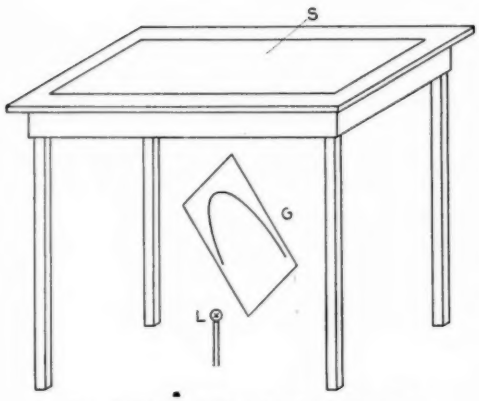


FIG. 3. Diagram of projection apparatus.

chromaticity, δE,

$$du=0,$$

hence $$dv=-dw,$$

and from Eq. (H-2b)

$$K\delta E=\sqrt{2}\,|dv|=\sqrt{2}\,|dw|.$$

If K is chosen so the average value of δE for the spectral colors is unity, and if, following Judd's notation, ΔE denotes the number of unit chromaticity steps between p_1 and p_2, then

$$K\Delta E=\sqrt{2}\,|v_2-v_1|=\sqrt{2}\,|w_2-w_1|.$$

The corresponding condition holds if both points lie on one of the lines in either of the other two logarithmic families.

In the general case where p_1 and p_2 do not lie along any one line of the logarithmic coordinate system it may be shown that

$$K^2\Delta E^2=(u_2-u_1)^2+(v_2-v_1)^2+(w_2-w_1)^2. \quad \text{(H-2c)}$$

It is clear, therefore, that along any of the lines of the logarithmic coordinate system the intervals between intersections correspond by hypothesis to a given fixed difference of chromaticity. In other directions the same difference of chromaticity is represented by intermediate distances. Thus the coordinate network constitutes a system of gauges whereby the lengths of the segments on the spectral curve may be visually compared to the consequences of Eq. (H-2b).

IV. Apparatus and Procedure

The actual arrangement of the apparatus is illustrated in Fig. 3. A table has a plate glass window built into its top. The upper surface of the glass is flush with the surrounding frame. Across the top is stretched a sheet of tracing cloth to serve as the translucent screen, S. On the cloth are drawn in black ink the lines of the special triangular coordinate system. The parent triangle of the system measures 33.33 inches on a side. The point-source, L, is a miniature incandescent electric lamp with concentrated coiled filament, mounted on an adjustable holder. The piece of plate glass, G, bearing the spectral curve is supported by a partially counterbalanced system of lever arms with ball-and-socket joints that allows the requisite six degrees of freedom and at the same time has sufficient frictional restraint so as to hold the plate in any position to which it is moved.

The following method was employed to inscribe the curve on the glass plate. Making use of experimental values of the spectral trichromatic coefficients, the spectral locus was plotted in the first place on coordinate paper. Rectangular coordinates were used, the data having been first transformed so that when plotted, the three spectral primaries R, G, B lay at the apices of an equilateral triangle. Segments corresponding to experimental values of $\delta\lambda$ were demarcated by pairs of short transverse lines at the proper intervals along the length of the curve. The curve was then traced over with black "Hektograph" ink and transferred by the "Hektograph" process to the glass plate, the surface of which had previously been coated with a thin transparent film of gelatine.

In order to locate the point source accurately at the origin of the three-dimensional coordinate system, the table top was first carefully leveled in two directions. A silk thread "plumb line" dropped through a pinhole at the center of the triangular system on the tracing-cloth (the glass in the table top was in two pieces) then provided a means to align the source with the coordinate system on the screen. As a matter of fact, the distance from source to screen does not affect

the results and may be adjusted to suit the convenience of the operator.

As already indicated, the experimental procedure consisted simply of moving the glass plate around in an attempt to find a position for which the segments on the projected curve were proportional in width to the interstices of the coordinate network at corresponding points on the screen. It was found possible by inspection to arrive at conditions very close to the most favorable. For the final adjustments the values of δE had to be calculated. In practice for the sake of greater precision the final adjustment was arrived at by stages. Once a setting had been made that approximated what was believed to be the best position of the spectral locus in the space reference frame, readings were taken and the spectral trichromatic coefficients were recalculated on that basis. From the latter, a new spectral curve was plotted and transferred to the glass plate as before. The apparatus was then ready to attempt a more accurate setting. Besides values of $\delta\lambda$ certain other experimental data relating to saturation sensibilities were also taken into account by calculative methods in reaching the final result.

With this simple apparatus it was possible to explore the field thoroughly with a minimum of time and effort. The results obtained with it agreed closely with parallel computations carried out from time to time. Numerical results, as stated, will appear in a paper to follow.

A Brightness Meter

MATTHEW LUCKIESH AND A. H. TAYLOR
Lighting Research Laboratory, General Electric Company, Nela Park, Cleveland, Ohio
(Received January 5, 1937)

THIS portable brightness meter was designed to measure a wide range of brightnesses. The practicable range of the meter is from about 0.025 to 50,000 footlamberts or 0.027 to 53,800 millilamberts or 0.00005 to 110 candles per square inch. The weight of the complete meter is 2 lb. 6 oz.

The meter has two eyepieces at the right-hand end. Through the upper eyepiece the observer views the magnified photometric field, consisting of two small silvered rectangles separated by a narrow clear strip, as shown in the sketch on the photograph. By means of the lens tube at the opposite end of the meter an image of the object to be measured is brought into focus in the plane of the photometric field, in the clear portion between and surrounding the two silvered rectangles. The latter reflect the image of a diffusing glass in front of the comparison lamp. The brightness of this glass is varied by means of a circular photographic film gradient and various filters of constant transmission factor may also be inserted. After a photometric balance is obtained the brightness of the test-object is read from a scale seen through the lower eyepiece at the right-hand end as shown on the photograph. The values are given both in candles per square inch and in footlamberts.

The scale of the meter reads from 0.0045 to 0.11 candle per square inch, also 2 to 50 footlamberts. Its range can be extended upwards by two multiplying screens with multiplying factors of 10 and 100, respectively. They can be used together with a multiplying factor of 1000. Brightnesses below the normal range of the meter may be read by means of a reducing screen with a factor 1/10, also by lowering the voltage of the comparison lamp. By use of the lower voltage and the reducing screen at the same time a factor of 1/100 is obtainable.

When any multiplying or reducing screen is in use the flat side of the button controlling it is faced towards the observer, with the multiplying factor plainly marked on it. The test object is clearly focused in the photometric field, hence there is no uncertainty as to the object being measured. The optical system is such that an object approximately 1 foot wide can be measured at a distance of about 500 feet.

When the comparison lamp voltage is set at the higher calibrated value, as it is for all but very low brightnesses, the color of the comparison field is approximately the same as that of the light from high wattage tungsten-filament lamps. This is accomplished by a blue filter fixed in position in the instrument.

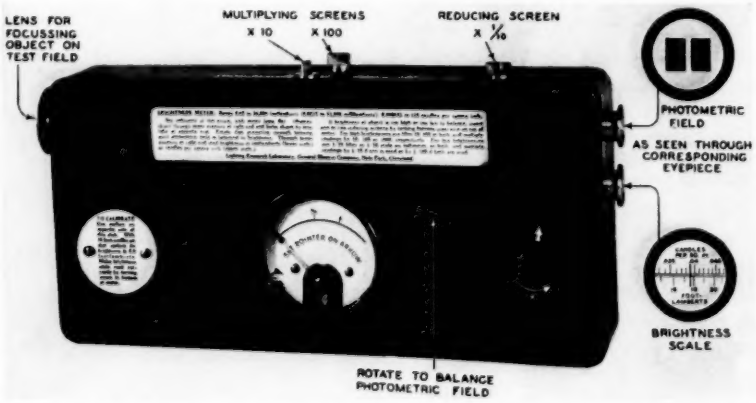


FIG. 1. A brightness meter of unusual range. The image of an object can be accurately focused upon the photometric field.

Hamilton's Characteristic Function and Bruns' Eiconal*

M. HERZBERGER
Kodak Research Laboratories, Rochester, New York

PROFESSOR SYNGE claims in his paper that I said unjustly that Bruns differed from Hamilton in his method, and that Bruns' method opened a new field for optical investigation. He claims that Bruns' function is identical with Hamilton's, and demonstrates that Hamilton's T function is very valuable for application to practical problems in optics.

I believe that I can demonstrate that there *is* a difference between the three characteristic functions, V, T, and W of Hamilton, and the three corresponding eiconals of Bruns, which we shall call $\tilde{S}$, $\tilde{W}$, and $\tilde{V}$; and that there are cases of great theoretical and practical importance where we can write $\tilde{S}$ and $\tilde{V}$ in explicit form, but not V and W. This fact, that we can obtain these functions explicitly in different cases, was very stimulating for the whole development of modern optics.

The case of T and $\tilde{W}$ is different; and that is the only thing that Professor Synge proves in his paper. Here both functions are identical, except that T is the homogeneous form of $\tilde{W}$ with two dummy variables. He proves that Hamilton solved most of the problems which later were attacked with the help of $\tilde{W}$; moreover, Hamilton himself has here made the step from the function of six variables to that of four variables, and has even suggested the introduction of the three variables a, b, c for the case of rotationally symmetrical systems. (Compare with my preceding paper.) I gladly acknowledge this fact, as well as the fact that Professor Synge has proved that some of the derivations are more elegant if the homogeneous function is used.

On the other hand, I have shown that there are cases when it is convenient or even necessary not to use T or $\tilde{W}$, for instance in telescopic systems. In many of these cases it is possible to get the Bruns eiconal in closed form, but not to obtain Hamilton's characteristic function.

Far be it from me to do anything to diminish Hamilton's glory. There was, and is, nobody in the science of geometrical optics whose work can compare in depth with his ideas. The characteristic function is one of his greatest inventions, and I should like to say, like Synge, that Bruns only rediscovered Hamilton's idea, without knowing it. With the historic part of Synge's paper I agree. However, I assert that the slight change in standpoint by going from the function of six variables to that of four variables has its significance. Bruns has found a great many results which could not be obtained without this change, and it seems to me that the whole development since Bruns would have been impossible without this change. I think that Bruns' paper is the most valuable contribution to optics since Hamilton's time. At the end of my reply I will defend my use of the word "eiconal."

Before I compare the two conceptions of Hamilton and Bruns, I wish to say a few words to make it understandable that such rediscoveries, like that of Bruns, are so much more frequent in optics than in any other mathematical or physical science. The reason is that geometrical optics is not yet a systematic science, in which each separate item is given its rightful place. Those who use it have generally not done more than create a fairly useful tool. That is, they have not searched for the general laws, but for approximations good enough to give the basis for lens design. The development of the general laws was in the hands of the mathematicians, who sometimes considered the problems of geometrical optics as a very interesting field in which to apply their pet mathematical theories. Hamilton used the calculus of variations, which he transformed into a very useful, and not yet sufficiently known, tool; Bruns, at the suggestion of Sophus Lie, applied the theory of contact transformations; the French mathematicians Dupin, Malus, Cauchy, and Gergonne used the theory of surfaces and differential line geometry; Möbius took special problems for applications of his theory of collineation; Gullstrand invented a very cumbersome tool, restricting himself to the use of elementary differential and integral cal-

* A reply to an article by Professor J. L. Synge, "Hamilton's Method in Geometrical Optics," J.O.S.A. **27**, 75 (1937).

culus, by creating differential invariants in differentiating the law of refraction; and so forth.

We may compare the usual scientific progress with the trip of an explorer, who has a map of the part of the earth which he wishes to explore, with the unknown regions indicated. If he finds some new island, he will give it a name, and the new island will be marked on all maps. The investigations of the past century in optics, however, can be compared with a crowd of explorers who climb a forested mountain of which they have no map, and where there is no real path. Since the explorers start from very different standpoints, they think that all they see is really unknown. Now it might easily happen that two of them come to the same place; one sees nothing special, the other climbs a tree and sees a wonderful view from the same point. Who has now discovered the point? We have not yet a map on which we can mark each discovery, but I very much hope that we shall soon be that far, and I believe that the ideas of Hamilton especially may pave the way to the summit, from which we can look over the whole field. But until that time, such rediscoveries will happen again and again, and no optical writer, not even Hamilton, is without fault. Other reasons are that we find three writers on optical problems using five different notations, and the papers are usually full of annoying calculations. In the case of Bruns and Hamilton, there is still another reason. Hamilton published his optical papers in the *Irish Transactions*. Even in my time there were in Germany only three copies of the *Irish Transactions* in all the public libraries. This was one of the reasons why I urged so much the translation and publication of Hamilton's papers, which Professor Prange edited in 1932. Now, since we have this translation, together with the excellent publication of Hamilton's papers by Synge and Conway, there should no longer be this excuse.

Now to the point in question. Let us discuss all the three characteristic functions one after another.

I. V AND $\bar{S}$

First, Hamilton's characteristic function $V(x, y, z, x', y', z')$ is defined as the optical path $\int_{xyz}^{x'y'z'} n ds$ from a point (x, y, z) in the object space to a point (x', y', z') in the image space. Of course, there is no optical system for which this function can be given explicitly. V is not a general function of its argument, but follows the two differential equations:

$$\left(\frac{\partial V}{\partial x}\right)^2+\left(\frac{\partial V}{\partial y}\right)^2+\left(\frac{\partial V}{\partial z}\right)^2=n^2,$$
$$\left(\frac{\partial V}{\partial x'}\right)^2+\left(\frac{\partial V}{\partial y'}\right)^2+\left(\frac{\partial V}{\partial z'}\right)^2=n'^2. \tag{1}$$

Bruns' eiconal $\bar{S}$ can be defined, according to Klein, as the optical path between the points of the two coordinate planes, so that we have

$$\bar{S}(x, y, x', y')=V(x, y, 0, x', y', 0). \tag{2}$$

In this way Synge seems actually justified, and the Bruns eiconal only a special case of Hamilton's characteristic function. But the essential point is that a ray is given by the coordinates of two of its points, so that the function $\bar{S}$ actually is sufficient to give all the qualities of the image formation; moreover, the function $\bar{S}$ is the general function of its variables, with the only restriction that

$$\frac{\partial^2 S}{\partial x \partial x'}\frac{\partial^2 S}{\partial y \partial y'}-\frac{\partial^2 S}{\partial x \partial y'}\frac{\partial^2 S}{\partial x' \partial y}\neq 0. \tag{3}$$

Moreover, we can in many interesting cases actually give $\bar{S}$ as a function of x, y, x', y' in explicit form, where the more general form V cannot be constructed. Someone may say that Hamilton's more general function has not been constructed, but may be in the future. I really doubt that Hamilton's function can be constructed, because the assumption that through initial and end-point goes one and only one ray can be fulfilled for the points of planes or curved surfaces, but not for the whole space. This does not mean that Hamilton's function is not very useful, and because of its generality sometimes more useful, but it can be applied only to problems in the small and not in the large, as can Bruns' functions.

I give two simple examples for an image formation in which Bruns' eiconal can be given in closed form:

$$\bar{S}=C(xx'+yy') \tag{4}$$

from my preceding paper (formula (5)).[1] We derive

$$n\xi = -Cx', \quad n'\xi' = Cx,$$
$$n\eta = -Cy', \quad n'\eta' = Cy. \tag{5}$$

We see that (4) gives the eiconal for a system in which both infinite planes have sharp images on the coordinate planes. We notice that the image is not orthoscopic. The above eiconal and the qualities of the corresponding image formation were investigated by Bruns. An eiconal which guarantees sharp image formation of the object plane alone is given by

$$\tilde{S} = f(x'y')x + g(x'y')y \tag{6}$$

with $$n\xi = -f(x'y'), \quad n\eta = -g(x'y'). \tag{7}$$

All these examples come from Bruns, but I could easily add others more complicated from recent papers.

II. W AND $\tilde{V}$

Now let us consider Hamilton's characteristic function W and the corresponding eiconal $\tilde{V}$. W is the optical path from the footpoint of the perpendicular dropped on the object ray from the origin to the point with the coordinates x', y', z'. It is a function of $n\xi$, $n\eta$, $n\zeta$, x', y', z'. Of course, $n\zeta$ could be eliminated with the help of

$$(n\xi)^2 + (n\eta)^2 + (n\zeta)^2 = n^2, \tag{8}$$

so that W is subject only to the single differential equation

$$\left(\frac{\partial W}{\partial x'}\right)^2 + \left(\frac{\partial W}{\partial y'}\right)^2 + \left(\frac{\partial W}{\partial z'}\right)^2 = n'^2. \tag{9}$$

We can define the corresponding Bruns eiconal by the equation

$$\tilde{V}(n\xi, n\eta, x', y') = W(n\xi, n\eta, (n^2 - (n\xi)^2 - (n\eta)^2)^{\frac{1}{2}}, x', y', 0) \tag{10}$$

and the same arguments would have to be repeated as for $\tilde{S}$ and V.

I give again some examples, where $\tilde{V}$ exists, but W cannot be constructed:

First, the corresponding eiconal

$$\tilde{V} = C(x'n\xi + y'n\eta). \tag{11}$$

We conclude from my preceding paper.[1]

$$x = Cx', \quad n'\xi' = Cn\xi,$$
$$y = Cy', \quad n'\eta' = Cn\eta. \tag{12}$$

We see that we have a telescopic system, which at the same times gives sharp images of the coordinate planes on each other. Special cases of formula (12) play a great role in modern investigation. $C=1$, for instance, gives the eiconal for refraction at a plane, $C=n'/n$ gives the only example of an image formation in which each point of the space has a sharp image, an image formation first investigated by Clerk Maxwell in a famous paper.

Let me give in this case a few more examples from my book *Strahlenoptik*.[2]

The eiconal for a sharp image of a plane on a curved surface is given by

$$\tilde{V} = -\beta' b - z'(n'^2 - c)^{\frac{1}{2}} + g(a). \tag{13}$$

In this formula is substituted for abbreviation

$$a = x^2 + y^2, \; b = xn'\xi' + yn'\eta', \; c = n'^2(\xi'^2 + \eta'^2) \tag{14}$$

and β', z', and g are three arbitrary functions of a whose geometric significance can easily be seen.

If we require that to each point shall correspond not a sharp image, but only a rotationally symmetrical image caustic, we get

$$\tilde{V} = -\rho b/2\sqrt{a} - z'(n'^2 - c)^{\frac{1}{2}} - n'\int_1^H \Delta(a, H)dH + g(a). \tag{15}$$

Here again, ρ and z' are functions of a alone, with significant geometrical meaning.

An important special case, where all the image caustics do not change their type is given by

$$\tilde{V} = -\rho b/2\sqrt{a} - z'(n'^2 - c)^{\frac{1}{2}} - (n'/k')\int_1^H \Delta(0, (H-1)k')dH. \tag{16}$$

These and many other investigations in the large could not have been made with Hamilton's functions of six variables.

[1] J. Opt. Soc. Am. **26**, 177–180 (1936).

[2] M. Herzberger: *Strahlenoptik* (J. Springer, 1931).

III. T AND $\tilde{W}$

It is different with Hamilton's T function and Bruns' angle eiconal $\tilde{W}$, the only functions Professor Synge has used in his article. Here the geometrical meanings are identical: both are the optical path between the footpoints of the perpendiculars dropped from the origins to object and image ray. T is a function of $n\xi$, $n\eta$, $n\zeta$, $n'\xi'$, $n'\eta'$, $n'\zeta'$, but since

$$\begin{aligned}(n\xi)^2+(n\eta)^2+(n\zeta)^2&=n^2,\\(n'\xi')+(n'\eta')+(n'\zeta')^2&=n'^2,\end{aligned}\tag{17}$$

T contains actually only four parameters. We have the equation

$$\begin{aligned}&T(n\xi,\ n\eta,\ (n^2-(n\xi)^2-(n\eta)^2)^{\frac{1}{2}},\ n\xi',\ n'\eta',\\&(n'^2-(n'\xi')^2-(n'\eta')^2)^{\frac{1}{2}}=\tilde{W}(n\xi,\ n\eta,\ n'\xi',\ n'\eta').\end{aligned}\tag{18}$$

In this case, there is actually not much difference between the two functions. Here I plead guilty. However, I take the opportunity to call attention to an error occurring frequently in papers using Hamilton's homogeneous function T. Since $n\xi$, $n\eta$, $n\zeta$ as well as $n'\xi'$, $n'\eta'$, $\eta'\zeta'$ are not independent coordinates, we cannot conclude from

$$\begin{aligned}dT=&\,xdn\xi+ydn\eta+zdn\xi\\&-x'dn'\xi'-y'dn'\eta'-z'dn'\zeta'\end{aligned}\tag{19}$$

the equations

$$\begin{aligned}\partial T/\partial n\xi&=x,\quad \partial T/\partial n'\xi'=x',\\\partial T/\partial n\eta&=y,\quad \partial T/\partial n'\eta'=y',\\\partial T/\partial n\zeta&=z,\quad \partial T/\partial n'\zeta'=z',\end{aligned}\tag{20}$$

because the derivation of a function of four variables gives not six independent functions. Let us take for example Synge's form for the characteristic function of the refraction at a sphere:

$$\begin{aligned}T=R[&(n'\xi'-n\xi)^2\\&+(n'\eta'-n\eta)^2+(n'\zeta'-n\zeta)^2]^{\frac{1}{2}}.\end{aligned}\tag{21}$$

We find

$$\begin{aligned}T\partial T/\partial n\xi&=-(n'\xi'-n\xi)R^2;\\T\partial T/\partial n\eta&=-(n'\eta'-n\eta)R^2;\\T\partial T/\partial n\zeta&=-(n'\zeta'-n\zeta)R^2;\end{aligned}\tag{22}$$

and therefore between $(\partial T/\partial n\xi)$, $(\partial T/\partial n\eta)$, $(\partial T/\partial n\zeta)$ the connection

$$(\partial T/\partial n\xi)^2+(\partial T/\partial n\eta)^2+(\partial T/\partial n\zeta)^2=R^2.\tag{23}$$

That means, if we regard Eqs. (20) as valid, xyz (and $x'y'z'$) would only determine the points of the refracting surface.

May I say at the end, that the same arguments hold true in the case of anisotropic and inhomogeneous media, because there, too, the manifold of extremals is always four-dimensional, that means, characterized by four parameters.

Let us recapitulate. There are many problems in which the $T(\tilde{W})$ function cannot be used. In all these cases the slight change of standpoint from Hamilton's characteristic function to Bruns' eiconal is a valuable step forward. In the case of the $T(\tilde{W})$ function it does not give any special advantage.

Finally, let me speak about the question, *characteristic function* or *eiconal*. I concede to Professor Synge that *characteristic function* is from a theoretical standpoint more significant. There was a commission in Germany to unify optical notation. Suggestions were invited from scientists in the field of optics outside Germany, and the nomenclature used in my paper was the proposal of the majority. The two kinds of functions have different regions of application; a distinction is therefore valuable. I believe the use of both names beside each other is quite in order, with express mention that Hamilton discovered the function, which Bruns only rediscovered with slight modification.

I intended at first to summarize here the contents of Bruns' paper, which I think is one of the best papers since Hamilton's time. Since this article has already attained undue length, I omit it, and recommend to the reader the article[3] in which I summarized the new results in his paper and compared them with its predecessors.

Note added on reading proof.—Professor Synge and I have exchanged copies of our manuscripts so I have had the opportunity of reading the following article. The editor does not wish this discussion to continue beyond the pages of this issue but allows me to make a few final remarks. I have looked over once more Hamilton's papers

[3] Zeits. f. Instrumentenkunde, **52**, 492 (1932).

and am still enthusiastic over the greatness of his ideas and the clearness with which he presents them.

Professor Synge succeeded in (5)* for one of my examples in deriving Hamilton's six-variable function. That shows that my remark before formula (11)* was too hasty.

Professor Synge is correct in that Hamilton always carefully determined the exact form of W and T by adding the condition that these functions are homogeneous of a certain order K in $\xi\eta\zeta$; ξ', $\eta'\zeta'$, respectively.

Hamilton uses with great virtuosity the factor K to simplify the calculations.

Professor Synge has not succeeded in giving Hamilton's V functions in the cases I mentioned. The difference between his function 3(1),* and the implicit function 3(2),* with the condition 3(3)* is plain enough. The remark of Synge at the end of his paragraph (3) is not understandable to me. Formula (5)* of this paper gives to each object ray the image ray, that means we know everything about the system. In his paragraph (4) Synge really gives only the four-variable function, which he calls boundary form of Hamilton's functions, which I called Bruns' eiconal.

I hold to my statement that in an isotropic media four parameters are sufficient to describe the optical laws.

The fact that Synge succeeded in giving the mixed characteristic in a case in which I believed it to be impossible, convinced me that I underestimated the applicability of Hamilton's theory. I am willing to accept the consequences and to adopt the name characteristics and the letters V, H, T used by Hamilton in my future work. Notwithstanding, I shall continue to use the four-variable functions instead of the six-variable functions, since they simplify in most cases the calculations considerably.

* J. L. Synge, "Hamilton's Characteristic Function and Bruns' Eiconal," J. Opt. Soc. Am. **27**, 138 (1937).

Hamilton's Characteristic Function and Bruns' Eiconal

J. L. SYNGE
University of Toronto, Toronto, Canada

I AM glad to have this opportunity to reply to the immediately preceding article by Dr. Herzberger under the above title, and to substantiate some of the statements made in my preceding article,[1] which was occasioned by, and in part critical of, an earlier article by Dr. Herzberger.[2]

1. THE QUESTION OF NOMENCLATURE

Dr. Herzberger rightly deplores the multiplicity of notations in geometrical optics, but does not seem disposed to approve any changes from the nomenclature of reference 2. It is agreed, however, that (without reference to trivial changes of sign) Bruns' point-eiconal $\bar{S}$ is a boundary value (for $z'=z=0$) of Hamilton's characteristic function V, Bruns' mixed eiconal $\bar{V}$ is a boundary value (for $z'=0$) of Hamilton's W, while Bruns' angle-eiconal $\bar{W}$ is precisely Hamilton's T. The subtle distinction which Dr. Herzberger draws between $\bar{W}$ and T (Eq. (18) of his reply) is of his own making.

There can be little doubt that this confusing duality of nomenclature arose through Bruns' complete ignorance of Hamilton's work in geometrical optics. Had he been acquainted with the basic literature of the subject he would hardly have invented entirely new names for concepts so immediately connected with those of Hamilton. In inventing the work "eikonal" and in believing that he was making a substantially new contribution, Bruns acted in good faith in the light of his knowledge, deficient as it was. The question for us is this: Are we justified, in the light of our present knowledge, in perpetuating Bruns' nomenclature?

First, there is the claim of long priority. Sometimes a crude early method is replaced by an important generalization, and in such cases there may be doubt as to whether the name of the originator or the improver should be perpetuated in connection with the method. But here no such question arises. Hamilton anticipated Bruns by sixty years with methods much more complete and general. The more different approaches there are to a difficult theory the better our understanding of it, but that is no reason for replacing the name of the originator by that of the author of a fresh approach. I shall try to indicate below in §2 what I conceive to be the slight distinction between the concepts of Hamilton and Bruns.

Secondly, there are considerations of the influence of nomenclature on the progress of science. It may not seem of much importance whether we use this word or that: actually, it has far-reaching effects on the course of the scientific stream. We are so conscious of progress in most branches of science in the last hundred years that we seldom turn to antique sources, except for their historical interest. We believe (in most cases rightly) that the work of the pioneers has been absorbed into the standard texts. If we write "eiconal" and "Bruns," we indicate to fresh workers in geometrical optics what their source of inspiration should be, and the history of the general theory of geometrical optics in the future is likely to involve a slow and partial rediscovery of results that Hamilton obtained long ago. If, on the other hand, we write "characteristic" and "Hamilton," we suggest the proper source of inspiration, and students of geometrical optics will be spared a repetition of the sorry commentary on their scholarship afforded by the incident of Bruns.

For the above reasons and in spite of the findings of the commission to which Dr. Herzberger refers, I beg respectfully to suggest to authoritative workers in geometrical optics that the time has come to drop the use of the expression "Bruns' eiconal" as standard nomenclature and to replace it by "Hamilton's characteristic," and further, since we have in Hamilton's papers such a comprehensive treatment of the basic theory of the subject, so carefully and consistently elaborated in respect of notation,

[1] "Hamilton's Method in Geometrical Optics," J. Opt. Soc. Am. **27**, 75 (1937). In the present paper I shall not use the subscript notation there proposed, but Hamilton's own notation, merely changing the signs of his W and T.
[2] J. Opt. Soc. Am. **26**, 177–180 (1936).

I suggest that Hamilton's notation should be made the basis of our modern notation: in particular, that the three leading characteristics should be known as V, W and T. We might well use some of the modern terminology, referring to V as the "point characteristic," W as the "mixed characteristic," and T as the "angle characteristic."

It may be of interest to note that three years before the paper of Bruns,[3] Thiesen[4] used the word "characteristic" to refer to the boundary-value of Hamilton's V, a function of four variables and in fact precisely the point-eiconal of Bruns.

2. Distinction Between the Points of View of Hamilton and Bruns

Hamilton considered what may be called the *direct* problem of geometrical optics: given an instrument, to find its optical behavior. This problem breaks into two: (1) given the instrument, to find a characteristic, (2) given a characteristic, to find the behavior. A complete discussion of these problems (including consideration of every conceivable type of instrument) cannot, of course, ever be given, but it may at least be said of Hamilton that he developed general methods, so that the solution of (1) was reduced to algebraic eliminations, while he developed a vast collection of results under (2), chiefly of a very general nature.

Bruns, on the other hand, was largely concerned with the *inverse* problem: given the behavior, to find the instrument. Actually this very important inverse problem breaks into two: (3) given the behavior, to find the characteristic, (4) given the characteristic, to find the instrument.

Dr. Herzberger has given examples of solutions of (3), and I shall deal with some of them below; had Bruns never written a word on geometrical optics, they could have been solved with the greatest ease on the basis of Hamilton's method. Any difference in method is quite trivial: Hamilton would have preferred to deal with the equations of the congruence of final rays in cases where Bruns would employ their intersections with a certain plane. Sometimes Hamilton's point of view gives added convenience, but the difference is so slight as hardly to warrant mention.

As for (4), the problem does not seem to have been touched, although it is the really important problem, without which the solution of (3) seems rather trivial. I do not know that Hamilton ever considered the inverse problem, except in a severely limited and practical way, as when he designed a doublet. It is hard to refrain from the suspicion that the simple forms of characteristics giving ideal imagery are too simple to be physically true, in the sense that no distribution of media could be found possessing such a simple characteristic. The actual characteristics that we know do not resemble them.

The difference between Hamilton and Bruns is largely a difference in mathematical psychology, as indeed Dr. Herzberger has indicated in his reply. Bruns preferred to base his method on the abstract theory of contact-transformations, but we must not forget that the ultimate source of contact-transformations is to be found in Hamilton's optical method, carried over by him into dynamics, and further developed by Jacobi.

It really seems that whatever distinction there is between the two points of view may best be discussed on the basis of Dr. Herzberger's statement above the inequality (3) of his reply, where he says of Bruns' $\bar{S}$, that it is "the general function of its variables," satisfying (3). In how wide a sense is the word "general" to be interpreted? Probably in a very wide sense indeed, if one considers the instrument merely as a mathematical machine performing a contact transformation. But if we think (as presumably we ought to do in geometrical optics) of media distributed in space, we must hesitate and admit our ignorance. Thus (if it is permissible to attempt a reconstruction of thought) I think that where Bruns might have said "The function $\bar{S}$ is arbitrary: let us choose it to satisfy certain requirements regarding imagery," Hamilton might have said "If there is an instrument satisfying certain conditions regarding imagery, it will have a characteristic and that characteristic will have a boundary value on a pair of planes. Let us see

[3] H. Bruns, K. sächs. Ges. d. wiss. Abhand. math.-phys. Cl. **21**, 323–436 (1895).

[4] M. Thiesen, Ann. d. Physik **45**, 821 (1892). In 1875 Maxwell had used the word in the same sense: *Scientific Papers of J. C. Maxwell*, Vol. 2 (Cambridge, 1890), p. 439.

what this boundary value must be in order that the conditions of imagery may be fulfilled." The distinction is slight—possibly nonexistent—and whatever results are obtained from the two points of view are identical.

We might be led to say that for homogeneous media Hamilton's T (Bruns' angle-eiconal) is a general function of its four arguments. But Hamilton showed[5] that in the case of a thin system of revolution certain relations between the aberration coefficients are always satisfied, although we have at our disposal an indefinite number of constants. How can we be certain that in the case of a thick instrument there are not certain identities, or more probably inequalities, due to the unavoidable fact that two media cannot occupy the same region in space?

3. The Construction of Hamilton's V from Boundary Values

Let me take up the challenges that Dr. Herzberger has thrown down. In the first place he states, after (3), that Bruns' function $\bar{S}$ can be given in explicit form where the more general form V cannot be constructed. What is meant by "explicit" and "constructed"? At first sight, we might say that "explicit" means expressible in terms of a finite number of the simple algebraic operations, but that would be a somewhat naive view for the twentieth century. If $y=\sin x$, is that explicit? If y is expressible as a Bessel function of x, is that explicit? If y is expressible by a quadrature in terms of x, is that explicit? I think the answer would be affirmative in every case (cf. Herzberger's Eq. (15)). The distinction between "explicit" and "implicit" is quite artificial, and while it is undoubtedly pleasant to have to deal with simple functions whose properties are well known, occasions arise from time to time in which new functions are defined in order to make implicit functions explicit, just as a Bessel function makes explicit the solution of a certain differential equation with certain boundary conditions.

Thus while Dr. Herzberger would be right if he said "simple explicit form" and "simply constructed," he is not justified in his statement if "explicit" and "constructed" are interpreted in a generous sense in conformity with modern mathematical ideas, and I also was wrong in my previous article[1] in accepting too readily the narrow definition of "construction" or "calculation."

To take as an example his expression (4), which may be written

$$(V)_{z'=z=0}=C(xx'+yy'), \tag{3.1}$$

the general characteristic V is constructed by eliminating x_0, x_0', y_0, y_0' from the five equations

$$\begin{gathered} V=C(x_0x_0'+y_0y_0')+n\rho+n'\rho', \\ Cx_0'+n(x_0-x)/\rho=0, \qquad Cy_0'+n(y_0-y)/\rho=0, \\ Cx_0+n'(x_0'-x')/\rho'=0, \quad Cy_0+n'(y_0'-y')/\rho'=0, \end{gathered} \tag{3.2}$$

where ρ, ρ' are contractions, namely,

$$\begin{aligned} \rho&=[(x_0-x)^2+(y_0-y)^2+z^2]^{\frac{1}{2}}, \\ \rho'&=[(x_0'-x')^2+(y_0'-y')^2+z'^2]^{\frac{1}{2}}. \end{aligned} \tag{3.3}$$

This process defines V in adequate manner, and if the point is made that it will be difficult to trace the properties of the system through the complicated algebraic processes, we may make the retort that without carrying them out at all we know just as much about the behavior of the system as we know from Dr. Herzberger's expression, Eq. (3.1) above.

4. V for a System Imaging Infinity on Two Planes

To show how unessential was the contribution of Bruns, let me indicate how Hamilton might have answered the question: What is value of V for points on the planes $z'=0$, $z=0$, if the infinite planes are sharply imaged on them? He might have given the answer as follows, not restricting himself to isotropic media as Bruns necessarily does. With initial components σ', τ' and final components σ, τ, we have in general

$$\begin{aligned} \sigma'&=-\partial V/\partial x', \quad \tau'=-\partial V/\partial y', \\ \sigma&=\partial V/\partial x, \qquad \tau=\partial V/\partial y. \end{aligned} \tag{4.1}$$

By the stated conditions, we know that, for $z'=z=0$, variations of x', y' leave σ', τ' unaltered and variations of x, y leave σ, τ unaltered. Thus, for $z'=z=0$,

[5] *The Mathematical Papers of Sir W. R. Hamilton*, Vol. 1 (Cambridge, 1931), p. 456.

$$\frac{\partial^2 V}{\partial x'^2}=\frac{\partial^2 V}{\partial x'\partial y'}=\frac{\partial^2 V}{\partial y'^2}=\frac{\partial^2 V}{\partial x^2}=\frac{\partial^2 V}{\partial x\partial y}=\frac{\partial^2 V}{\partial y^2}=0, \quad (4.2)$$

and hence, with A, B, C, D, E arbitrary constants,

$$(V)_{z'=z=0}=(Ax+By+C)(Dx'+Ey'+1), \quad (4.3)$$

of which (3.1) is a special case.

As for a sharp image of infinity on the plane $z=0$ alone, we have in that case only the first three of (4.2) and hence

$$\begin{gathered}(V)_{z'=z=0}=f(x, y)x'+g(x, y)y'+h(x, y),\\ \sigma'=-f(x, y), \quad \tau'=-g(x, y),\end{gathered} \quad (4.4)$$

of which Dr. Herzberger's (6) is a special case.

5. Hamilton's W and Bruns' $\tilde{V}$

For comparison between W and $\tilde{V}$, Dr. Herzberger writes in his Eq. (10) a form for Hamilton's W which is not to be found in the works of Hamilton; it is, as a matter of fact, a meaningless form. Hamilton's W is a function of the final components σ, τ, υ and the initial coordinates x', y', z'. When the final medium is homogeneous, the form of W is partially indeterminate, and we may remove this indeterminacy by one of two devices. Either we can treat σ, τ as independent variables (υ being given in terms of them by $\upsilon^2=n^2-\sigma^2-\tau^2$, if the medium is isotropic), or we can make W homogeneous (say of degree unity) in σ, τ, υ. Thus we have for W either the form $W(x', y', z', \sigma, \tau)$ or $W(x', y', z', \sigma, \tau, \upsilon)$ with the identity[6]

$$\sigma\frac{\partial W}{\partial \sigma}+\tau\frac{\partial W}{\partial \tau}+\upsilon\frac{\partial W}{\partial \upsilon}=W, \quad (5.1)$$

but never the meaningless form (10) of Dr. Herzberger which would give no definite meaning to $\partial W/\partial\sigma$ (or $\partial W/\partial(n\xi)$ in his notation).

The true relation between Hamilton's W and Bruns' $\tilde{V}$ is simply

$$\tilde{V}(x', y', \sigma, \tau)=[W(x', y', z', \sigma, \tau)]_{z'=0}; \quad (5.2)$$

in fact, $\tilde{V}$ is a boundary value of W. (I have changed the sign of Hamilton's W.)

[6] Reference 5, p. 183.

6. The Construction of W

Opinions may differ regarding the practicability of making use of a V function defined by (3.2), but in the case of W, Dr. Herzberger makes, above his Eq. (11), a false statement when he says: "I give again some examples, where $\tilde{V}$ exists, but W cannot be constructed." W can be constructed quite simply, and I have given it explicitly in the case of refraction for several years in my lectures. The derivation is as follows, starting from Dr. Herzberger's (11), which in the less clumsy notation[7] of Hamilton reads

$$[W(x', y', z', \sigma, \tau)]_{z'=0}=C(x'\sigma+y'\tau). \quad (6.1)$$

In accordance with the rules given by Hamilton, W is given by the elimination of x_0', y_0' from the equations

$$\begin{gathered}W=C(x_0'\sigma+y_0'\tau)+n'\rho',\\ C\sigma+n'(x_0'-x')/\rho'=0,\\ C\tau+n'(y_0'-y')/\rho'=0,\end{gathered} \quad (6.2)$$

where $$\rho'=[(x_0'-x')^2+(y_0'-y')^2+z'^2]^{\frac{1}{2}}. \quad (6.3)$$

From (6.3) and the last two of (6.2) we have at once

$$\begin{gathered}C^2(\sigma^2+\tau^2)=n'^2(1-z'^2/\rho'^2),\\ \rho'=\epsilon n'z'[n'^2-C^2(\sigma^2+\tau^2)]^{-\frac{1}{2}},\end{gathered} \quad (6.4)$$

where $\epsilon=\pm1$, and we have also from (6.2)

$$x_0'=x'-C\sigma\rho'/n', \quad y_0'=y'-C\tau\rho'/n', \quad (6.5)$$

so that by the first of (6.2)

$$\begin{aligned}W&=C(x'\sigma+y'\tau)-C^2(\sigma^2+\tau^2)\rho'/n'+n'\rho'\\ &=C(x'\sigma+y'\tau)+\rho'[n'^2-C^2(\sigma^2+\tau^2)]/n',\end{aligned} \quad (6.6)$$

and so by (6.4) we have the simple explicit form of W (contrary to Dr. Herzberger's assertion)

$$W=C(x'\sigma+y'\tau)+\epsilon z'[n'^2-C^2(\sigma^2+\tau^2)]^{\frac{1}{2}}. \quad (6.7)$$

We are to choose $\epsilon=-1$ or $+1$ according as the ray through x', y', z' travels in the positive or negative sense of the z' axis, since $\upsilon'=-\partial W/\partial z'$.

[7] To use a double symbol, such as $n\xi$, for an independent variable with respect to which we have to differentiate, is extremely clumsy.

7. W for a System Imaging a Plane on a Given Surface

Dr. Herzberger quotes certain results given in his book,[8] and concludes: "These and many other investigations in the large could not have been made with Hamilton's functions of six variables." This statement is misleading, because for these problems Hamilton could certainly have used his W, a function of five variables. It is also false, because, without the use of any methods outside those published by Hamilton a century ago we can attack these problems with great ease, generality and elegance.

I shall use accents for initial quantities, and change the sign of Hamilton's W, so that the equations of the final rays are

$$x-z\sigma/v=-\partial W/\partial\sigma, \quad y-z\tau/v=-\partial W/\partial\tau, \qquad W=W(x', y', z', \sigma, \tau), \tag{7.1}$$

when the final medium is homogeneous and isotropic. The isotropy is not however used in the work immediately following.

Generalizing the problem of Dr. Herzberger, let us find $(W)_{z'=0}$ for a system (not necessarily of revolution) which gives a precise image of the plane $z'=0$ on a given surface, with assigned point-to-point imagery, which may be defined by writing

$$x=f(x', y'), \quad y=g(x', y'), \quad z=h(x', y'), \tag{7.2}$$

f, g, h being given functions. The initial medium is quite general, but the final medium will be supposed to be homogeneous, though not necessarily isotropic.

Now generally

$$W=V-\sigma x-\tau y-vz. \tag{7.3}$$

For any ray through $(x', y', 0)$ we may take the final point (x, y, z) at the position given by (7.2). V is independent of σ, τ, and *hence we have as the most general form of the required characteristic*

$$(W)_{z'=0}=F(x', y')-\sigma f(x', y')-\tau g(x', y')-vh(x', y'), \tag{7.4}$$

where F is an undetermined function and v a given function of σ, τ defined by the properties of the final medium, being

$$v=(n^2-\sigma^2-\tau^2)^{\frac{1}{2}} \tag{7.5}$$

when the medium is isotropic and the final rays proceed in the positive sense of the z axis. The formula (7.4) includes Herzberger's result (Eq. (4) of his book, p. 167) as a very special case, obtained by using the conditions of rotational symmetry, and introducing the notation

$$F(x', y')=g(a), \quad f(x', y')/x'=g(x', y')/y'=\tfrac{1}{2}\beta_s'(a), \quad h(x', y')=z^{*\prime}(a), \quad x'^2+y'^2=a, \quad x'\sigma+y'\tau=b, \quad \sigma^2+\tau^2=c. \tag{7.6}$$

It is perhaps of interest to remark that I did not actually arrive at (7.4) by the argument given above: I integrated (7.1), and then saw from the simplicity of the result that there must be a simpler argument. This is an important point, because the Eqs. (7.1) are emphatically equations of Hamilton, not of Bruns. Bruns always suppressed the z terms by considering only intersections of the rays with the plane $z=0$. It is extremely doubtful whether one could proceed from Bruns' equations to (7.4) without making use of clumsy and unnatural methods.

Dr. Herzberger's statement that Hamilton's function "can be applied only to problems in the small and not in the large, as can Bruns' functions" is completely false, if it means anything at all. Seduced by the attractive properties of conjugate planes in Gaussian optics, Bruns attached altogether too much importance to the planes $z'=0$, $z=0$, as the above result indicates clearly.

8. Hamilton's Method for Caustics

As regards the results (15) and (16) of Dr. Herzberger's reply, I have not had time to do more than see that their proof is based on the use of Hamilton's W function, the modification being trivial. It is evident, however, that these results are developed in complete ignorance of the general and powerful methods which Hamilton employed for the determination of caustic surfaces[9] by means of his functions W (for object

[8] M. Herzberger, *Strahlenoptik* (Springer, 1931), pp. 167 *et seq.*

[9] Reference 5, p. 265.

at a finite distance) and S (for object at infinity). The fact that Hamilton developed his theory to cover cases more general than those contemplated in the usual theory of optical instruments is no excuse for the complete neglect of these general methods, from which the applicable forms can easily be deduced. Under the assumption (apparently erroneous) that the leading workers in geometrical optics were acquainted with the most important literature of their subject, and were capable of deducing the particular from the general, I have not published the simplified form of Hamilton theory, although I have given it in lectures in the University of Toronto for a number of years. It may be well to put it down here, although an immediate and obvious deduction from the Eqs. (7.1) of the final rays.

Consider the congruence of rays from an initial point x', y', z', the components of a final ray being σ, τ, v. Confining our attention to a final medium which is homogeneous and isotropic, we have

$$v^2=n^2-\sigma^2-\tau^2, \quad v\delta v=-\sigma\delta\sigma-\tau\delta\tau. \quad (8.1)$$

If two rays (σ, τ), $(\sigma+\delta\sigma, \tau+\delta\tau)$ from (x', y', z') intersect to the first order at x, y, z, then (7.1) hold when we substitute $(\sigma+\delta\sigma, \tau+\delta\tau)$ for (σ, τ). Hence, denoting partial derivatives by subscripts, we have

$$\begin{aligned}&[W_{\sigma\sigma}-z(v^2+\sigma^2)/v^3]\delta\sigma+[W_{\sigma\tau}-z\sigma\tau/v^3]\delta\tau=0,\\ &[W_{\sigma\tau}-z\sigma\tau/v^3]\delta\sigma+[W_{\tau\tau}-z(v^2+\tau^2)/v^3]\delta\tau=0.\end{aligned} \quad (8.2)$$

Eliminating the differentials we have

$$[z(v^2+\sigma^2)-W_{\sigma\sigma}v^3][z(v^2+\tau^2)-W_{\tau\tau}v^3]-[z\sigma\tau-W_{\sigma\tau}v^3]^2=0, \quad (8.3)$$

a quadratic equation for the two values of z corresponding to the foci on the final ray σ, τ.

The Cartesian equation of the caustic surface due to a source of light at (x', y', z') is to be found by eliminating σ, τ from (8.3) and the two equations (7.1), that is,

$$x-z\sigma/v=-W_\sigma, \quad y-z\tau/v=-W_\tau. \quad (8.4)$$

This result contains the whole theory of caustic surfaces from a source at finite distance, the final medium being homogeneous and isotropic.

If the optical system is of revolution, then W involves its arguments only in the forms

$$z', \epsilon_0=x'^2+y'^2, \quad \epsilon_1=x'\sigma+y'\tau, \quad \epsilon_2=\sigma^2+\tau^2. \quad (8.5)$$

Indicating partial derivatives with respect to ϵ_0, ϵ_1, ϵ_2 by numerical subscripts, we have

$$\begin{aligned}&W_\sigma=W_1x'+2W_2\sigma, \quad W_\tau=W_1y'+2W_2\tau,\\ &W_{\sigma\sigma}=W_{11}x'^2+4W_{12}x'\sigma+4W_{22}\sigma^2+2W_2,\\ &W_{\sigma\tau}=W_{11}x'y'+2W_{12}(x'\tau+y'\sigma)+4W_{22}\sigma\tau,\\ &W_{\tau\tau}=W_{11}y'^2+4W_{12}y'\tau+4W_{22}\tau^2+2W_2.\end{aligned} \quad (8.6)$$

Hence, choosing axes so that $y'=0$, *the caustic surface formed in a system of revolution by a point source at* $(x', 0, z')$ *is given (as an equation connecting* x, y, z*) by eliminating* σ, τ *from (8.3) and (8.4), in which we are now to substitute*

$$\begin{aligned}&W_\sigma=W_1x'+2W_2\sigma, \quad W_\tau=2W_2\tau,\\ &W_{\sigma\sigma}=W_{11}x'^2+4W_{12}x'\sigma+4W_{22}\sigma^2+2W_2,\\ &W_{\sigma\tau}=2W_{12}x'\tau+4W_{22}\sigma\tau,\\ &W_{\tau\tau}=4W_{22}\tau^2+2W_2.\end{aligned} \quad (8.7)$$

This result contains the whole theory of caustic surfaces from a source at finite distance for an instrument of revolution, the final medium being homogeneous and isotropic.

9. Errors in Regard to Hamilton's T

Dr. Herzberger now admits that "there is not much difference" between Hamilton's T and Bruns' $\bar{W}$ (his Eq. (18)). The only apparent difference there is arises from the peculiar form which he has thought fit to give the left-hand side of his equation, a form which Hamilton never employed. From the lines following Eq. (18) I conclude that he does not understand the two devices suggested by Hamilton, *viz.* (*a*) to regard T as a function of four independent variables, (*b*) to construct a homogeneous function T of six variables. If he had read Hamilton's very clear statement with regard to this[10] he would have saved himself from this confusion, and would not have gone to the trouble of

[10] Reference 5, pp. 182, 183.

exploding (in his Eqs. (20) to (22)) a quite puerile error into which nobody with any claim to competency in this field could possibly fall—a "howler," in fact.

Having exploded this fictitious error, Dr. Herzberger makes one himself, if I am correct in understanding that by the words "the same arguments hold true in the case of anisotropic and inhomogeneous media" he wishes to convey that his Eqs. (20) do not hold in the case of these more general media. If he had gone to the trouble of consulting Hamilton's work, he would have found his equations (20) correctly stated as valid for heterogeneous media.[11]

10. Conclusion

Dr. Herzberger has produced no evidence whatever that there exists a distinct "method of Bruns," except insofar as the employment of boundary values of Hamilton's V and W may be considered a "method;" as mentioned,[4] this idea (for V) was already familiar to Maxwell and Thiesen before Bruns published. There is nothing which can be done "by the method of Bruns" which cannot be done as easily (and sometimes more easily and more generally, as in (7.4)) by the method of Hamilton.

The time has come to demand from those who claim authority in the theory of geometrical optics a standard of scholarship comparable to that obtaining in other branches of applied mathematics. The first criterion of scholarship is a reasonable familiarity with the most important literature, and there can be no doubt that Hamilton's contribution to the theory of geometrical optics remains a mine of general methods, almost completely unstudied. In few subjects has a great mathematician given such a detailed and general development of fundamental theory, and in no other subject has the work of a master been rejected as inapplicable because the students of the subject would not take the trouble to overcome the difficulties of understanding a general method and developing its applications.

[11] Reference 5, p. 176 (Eqs. (L′)). He would also have found them in my article, reference 1, Eqs. (3.3).

Optical Society of America, Inc.

Notice of Election of Officers

In accordance with Article III, Section 6, of the By-Laws of the Optical Society of America, the following officers and directors are to be elected at or before the Annual Meeting in October, 1937:

President, Vice President — for a term of two years

Secretary, Treasurer, 3 Members at large of the Board of Directors — for a term of four years.

The present officers are:

	Term Expires at Annual Meeting in:
President—A. C. Hardy	1937
Vice President—R. C. Gibbs	1937
Secretary—L. B. Tuckerman	1937
Treasurer—Henry F. Kurtz	1937
Past President—W. B. Rayton	1937
Editor—F. K. Richtmyer (elected by Directors)	1937
Directors at Large:	
A. H. Compton	
H. E. Ives	1937
C. W. Keuffel	
K. S. Gibson	
G. R. Harrison	1939

In accordance with the By-Laws, Article III, Section 4:

> "No member at large of the board of directors shall be eligible for reelection."

A. H. Compton, H. E. Ives and C. W. Keuffel are not eligible for reelection.

The following Committee on Nominations has been appointed and confirmed in accordance with Article III, Section 6, of the By-Laws:

W. B. Rayton, *Chairman*
Bausch & Lomb Optical Company,
Rochester, N. Y.

Walter Clark,
Research Laboratory,
Eastman Kodak Company,
Rochester, N. Y.

A. W. Kenney,
E. I. DuPont de Nemours & Company,
Experimental Station,
Wilmington, Delaware.

G. B. Pegram,
Columbia University,
New York, N. Y.

A. G. Worthing,
Department of Physics,
University of Pittsburgh,
Pittsburgh, Pa.

It will be the duty of this committee to make nominations for the above-mentioned election. The Committee report is to be made not later than July 15, 1937.

Although, after the committee has made its report, an opportunity will be given for additional nominations by petition, members are urged, as soon as possible, to suggest names which should be considered by the committee. Suggestions may be sent to the Chairman or any other member of the Committee.

Only regular and honorary members are eligible to election. A list of members and past officers was published in the J. O. S. A. Vol. 25, No. 4, pp. 103–124, April 1935. Later transfers to regular membership were given in the Secretary's *Annual Report* published in the J. O. S. A. Vol. 26, No. 7, p. 299, July 1936, and in the Minutes of the Eighth Annual Meeting of the Directors of the Optical Society of America, Incorporated, held October 28, 1936, published in the J. O. S. A. Vol. 27, No. 1, page 57, January 1937.

L. B. Tuckerman, *Secretary.*